box SEQ A

Express Mail Label No. EL280661249US

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Attorney Docket No.	STK-077
First Named Inventor	Hermann Oppermann et al.
Title	MODIFIED PROTEINS AND DNAs OF THE TGF-β SUPERFAMILY, INCLUDING MODIFIED MORPHOGENIC PROTEINS

APPLICATION ELEMENTS	ADDRESS TO: **Box Patent Application** **Assistant Commissioner for Patents** **Washington, D.C. 20231**
1. ☒ Fee Transmittal Form	**ACCOMPANYING APPLICATION PARTS**
2. ☒ Specification [Total Pages - 143] - Written Description - (88 pages) - Claims - (3 pages) - Abstract - (1 page) - Paper Copy of Sequence List - (40 page) - Sheets of Drawings - (11 sheets) ☐ Formal ☒ Informal - Paper copy of Sequence Submission - (40 pgs.)	7. ☐ 37 CFR 3.73(b) Statement (when there is an assignee) ☐ Power of Attorney 8. ☐ English Translation Document *(if applicable)* 9. ☐ Information Disclosure Statement (IDS)/PTO-1449 ☐ Copies of IDS Citations
3. ☐ Oath or Declaration [Total Pages] a. ☐ Newly executed (original) b. ☐ Copy from a prior application (37 CFR 1.63(d)) *(for continuation/divisional with Box 17 completed)* ***[Note Box 4 below]***	10. ☐ Preliminary Amendment ☐ Drawings [Total Sheets] ☐ Letter to Official Draftsperson Including Drawings [Total Pages] 11. ☒ Return Receipt Postcard
4. ☐ Incorporation by Reference (usable if Box 3b is checked) The entire Disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 3b, is considered as being part of the disclosure of the accompanying application and is hereby incorporated by reference therein.	12. ☐ Small Entity Statement(s) ☐ Statements filed in prior application, (Status still proper and desired) 13. ☐ Certified Copy of Priority Document(s)
5. ☐ Microfiche Computer Program *(Appendix)*	14. ☐ Deletion of Inventor(s) Signed statement attached deleting inventor(s) named in the prior application.
6. ☒ Nucleotide and/or Amino Acid Sequence Submission ☒ Computer Readable Copy ☒ Paper Copy (identical to computer copy) ☒ Statement verifying identify of above copies	15. ☒ Patent Application Data Entry Form 16. ☐ Other:

17. ☐ **If a CONTINUING APPLICATION,** check appropriate box and supply the requisite information:
☐ Continuation ☐ Divisional ☐ Continuation-in-part (CIP) of prior application Serial No. ____/____.
Priority to the above application(s) is claimed under 35 U.S.C. 120.
Prior application information: Examiner: _____. Group/Art Unit: _____.

18. ☒ **Priority - 35 U.S.C. 119**
☒ Priority of application Serial No. 60/103,418 filed on October 7, 1998 is claimed under 35 U.S.C. 119.
☐ The certified copy has been filed in prior U.S. application Serial No. ____/____ on ____.
☐ The certified copy will follow.

CORRESPONDENCE ADDRESS	SIGNATURE BLOCK
Direct all correspondence to: Patent Administrator Testa, Hurwitz & Thibeault, LLP High Street Tower 125 High Street Boston, MA 02110 Tel. No.: (617) 248-7000 Fax No.: (617) 248-7100	Respectfully submitted, Date: August 16, 1999 Reg. No. 40,702 Tel. No.: (617) 248-7045 Fax No.: (617) 248-7100 Timothy P. Linkkila Attorney for Applicants Testa, Hurwitz & Thibeault, LLP High Street Tower 125 High Street Boston, MA 02110

VITO7557/47.A814129_1

JC688 U.S. PTO 08/16/99
JC553 U.S. PTO 08/16/99

Inventor Information

Inventor One Given Name :: Hermann
Family Name :: Oppermann
Name Suffix ::
Postal Address Line One :: 22 Summer Hill
Postal Address Line Two ::
City :: Medway
State/Province :: Massachusetts
Country :: USA
Postal or Zip Code ::
City of Residence :: Medway
State/Prov. of Residence :: Massachusetts
Country of Residence :: USA
Citizenship :: USA

Inventor Two Given Name :: Mei-Sheng
Family Name :: Tai
Name Suffix ::
Postal Address Line One :: 47 Monroe Street
Postal Address Line Two ::
City :: Shrewsbury
State/Province :: Massachusetts
Country :: USA
Postal or Zip Code ::
City of Residence :: Shrewsbury
State/Prov. of Residence :: Massachusetts
Country of Residence :: USA
Citizenship :: USA

Inventor Three Given Name :: John
Family Name :: McCartney
Name Suffix ::
Postal Address Line One :: 210 Mellen Street
Postal Address Line Two ::
City :: Holliston
State/Province :: Massachusetts
Country :: USA
Postal or Zip Code ::
City of Residence :: Holliston
State/Prov. of Residence :: Massachusetts
Country of Residence :: USA
Citizenship :: USA

Correspondence Information

Correspondence Customer Number :: 021323

Application Information

Title Line One :: MODIFIED PROTEINS AND DNAs OF THE TGF-β SUPERFAMILY,
Title Line Two :: INCLUDING MODIFIED MORPHOGENIC PROTEINS
Total Drawing Sheets :: 11
Formal Drawings ::
Application Type :: Utility
Docket Number :: STK-077
Licensed - U S Government Agency ::
Contract Number ::
Grant Number ::
Secrecy Order in Parent Application ::

Representative Information

Representative Customer Number :: 021323

Continuity Information

This application is a ::
>Application One ::
Filing Date ::
Patent Number ::
which is a ::
>>Application Two ::
Filing Date ::
Patent Number ::

Prior Foreign Applications

Foreign Application One ::
Filing Date ::
Country ::
Priority Claimed ::

VITO7557/47.A814177_1

Express Mail Label No. : EL280661249US

APPLICATION

FOR

UNITED STATES

LETTERS PATENT

SPECIFICATION

(For Attorney Docket No. STK-077)

TO ALL WHOM IT MAY CONCERN:

Be it known that we, **Hermann Oppermann, Mei-Sheng Tai and John McCartney,** all citizens of the United States of America, and residing at 22 Summer Hill, Medway, 47 Monroe Street, Shrewsbury, and 210 Mellen Street, Holliston, respectively, in the Commonwealth of Massachusetts, in the United States of America, have invented new and useful improvements in

MODIFIED PROTEINS AND DNAs OF THE TGF-β SUPERFAMILY, INCLUDING MODIFIED MORPHOGENIC PROTEINS

of which the following is a specification.

MODIFIED PROTEINS AND DNAs OF THE TGF-β SUPERFAMILY, INCLUDING MORPHOGENIC PROTEINS

Continuing Application Data

The instant utility application is based on provisional patent application 60/103,418 filed on October 7, 1998, the entire contents of which is herein incorporated by reference; and, the instant application is related to co-pending utility applications U.S.S.N. _____ and _____ (Attorney Docket Nos. STK-075 and STK-076) filed on even date herewith, also based on the aforementioned provisional application, the disclosures of which are herein incorporated by reference.

Field of the Invention

The invention relates to modified proteins, and DNAs encoding the same, which are biosynthetic, chemosynthetic or recombinant constructs derived from the TGF-β superfamily of structurally-related proteins, including modified morphogenic proteins.

Background of the Invention

The TGF-β superfamily includes five distinct forms of TGF-β (Sporn and Roberts (1990) in Peptide Growth Factors and Their Receptors, Sporn and Roberts, eds., Springer-Verlag: Berlin pp. 419-472), as well as the differentiation factors Vg-1 (Weeks and Melton (1987) Cell 51: 861-867), DPP-C polypeptide (Padgett et al. (1987) Nature 325: 81-84), the hormones activin and inhibin (Mason et al. (1985) Nature 318: 659-663; Mason et al. (1987) Growth Factors 1: 77-88), the Mullerian-inhibiting substance, MIS (Cate et al. (1986) Cell 45:685-698), osteogenic and morphogenic proteins OP-1 (PCT/US90/05903), OP-2 (PCT/US91/07654), OP-3 (PCT/WO94/10202), the BMPs, (see U.S. Patent Nos. 4,877,864; 5,141,905; 5,013,649; 5,116,738; 5,108,922; 5,106,748; and 5,155,058), the developmentally regulated protein Vgr-1 (Lyons et al. (1989) Proc. Natl. Acad. Sci. USA 86: 4554-4558) and the growth/differentiation

factors GDF-1, GDF-3, GDF-9 and dorsalin-1 (McPherron et al. (1993) J. Biol. Chem. 268: 3444-3449; Basler et al. (1993) Cell 73: 687-702), to name but a few.

The proteins of the TGF-β superfamily are disulfide-linked homo- or heterodimers that are expressed as large precursor polypeptide chains containing a hydrophobic signal sequence, a long and relatively poorly conserved N-terminal pro region sequence of several hundred amino acids, a cleavage site, a mature domain comprising an N-terminal region that varies among the family members and a highly conserved C-terminal region. This C-terminal region, present in the processed mature proteins of all known family members, contains approximately 100 amino acids with a characteristic cysteine motif having a conserved six or seven cysteine skeleton. Although the position of the cleavage site between the mature and pro regions varies among the family members, the cysteine pattern of the C-terminus of all of the proteins is in the identical format, ending in the sequence Cys-X-Cys-X (Sporn and Roberts (1990), supra).

A unifying feature of the biology of the proteins of the TGF-β superfamily is their ability to regulate developmental processes, including endochondral bone morphogenesis. These structurally related proteins have been identified as being involved in a variety of developmental events. For example, TGF-β and the polypeptides of the inhibin/activin group appear to play a role in the regulation of cell growth and differentiation. MIS causes regression of the Mullerian duct in development of the mammalian male embryo, and dpp, the gene product of the Drosophila decapentaplegic complex, is required for appropriate dorsal-ventral specification. Similarly, Vg-1 is involved in mesoderm induction in Xenopus, and Vgr-1 has been identified in a variety of developing murine tissues. Regarding bone formation, proteins in the TGF-β superfamily, for example OP-1 and a subset of other proteins identified as BMPs (bone morphogenic proteins) play the major role. OP-1 (BMP-7) and other osteogenic proteins have been produced using recombinant techniques (U.S. Patent No. 5,011,691 and PCT Application No. US 90/05903) and shown to be able to induce formation of true endochondral bone *in vivo*. The osteogenic proteins generally are classified in the art as a subgroup of the TGF-β superfamily of growth factors (Hogan (1996), Genes & Development, 10:1580-1594).

Recently certain members of this same family of proteins have been recognized to be morphogenic, i.e., capable of inducing the developmental cascade of tissue morphogenesis in a

mature mammal (See PCT Application No. US 92/01968). In particular, these morphogens are capable of inducing the proliferation of uncommitted progenitor cells, and inducing the differentiation of these stimulated progenitor cells in a tissue-specific manner under appropriate environmental conditions. In addition, the morphogens are capable of supporting the growth and maintenance of these differentiated cells. These morphogenic activities allow the proteins to initiate and maintain the developmental cascade of tissue morphogenesis in an appropriate, morphogenically permissive environment, stimulating stem cells to proliferate and differentiate in a tissue-specific manner, and inducing the progression of events that culminate in new tissue formation. These morphogenic activities also allow the proteins to induce the "redifferentiation" of cells previously stimulated to stray from their differentiation path. Under appropriate environmental conditions it is anticipated that these morphogens also may stimulate the "redifferentiation" of committed cells.

Members of this morphogenic class of proteins include the mammalian osteogenic protein-1 (OP-1, also known as BMP-7, and the Drosophila homolog 60A), osteogenic protein-2 (OP-2, also known as BMP-8), osteogenic protein-3 (OP-3), BMP-2 (also known as BMP-2A or CBMP-2A, and the *Drosophila* homolog DPP), BMP-3, BMP-4 (also known as BMP-2B or CBMP-2B), BMP-5, BMP-6 and its murine homolog Vgr-1, BMP-9, BMP-10, BMP-11, BMP-12, GDF3 (also known as Vgr2), GDF-8, GDF-9, GDF-10, GDF-11, GDF-12, BMP-13, BMP-14, BMP-15, GDF-5 (also known as CDMP-1 or MP52), GDF-6 (also known as CDMP-2 or BMP-13), GDF-7 (also known as CDMP-3 or BMP-12), the *Xenopus* homolog Vgl and NODAL, UNIVIN, SCREW, ADMP, and NEURAL, to name but a few.

By way of illustration using exemplary family members, the tertiary and quaternary structure of both TGF-β2 and OP-1 have been determined. Although TGF-β2 and OP-1 exhibit only about 35% amino acid identity in their respective amino acid sequences the tertiary and quaternary structures of both molecules are strikingly similar. Both TGF-β2 and OP-1 are dimeric in nature and have a unique folding pattern involving six of the seven C-terminal cysteine residues. In each subunit four cysteines bond to generate an eight residue ring, and two additional cysteine residues form a disulfide bond that passes through the ring to form a knot-like structure. With a numbering scheme beginning with the most N-terminal cysteine of the 7 conserved cysteine residues assigned number 1, the 2nd and 6th cysteine residues bond to close

one side of the eight residue ring while the 3rd and 7th cysteine residues close the other side. The 1st and 5th conserved cysteine residues bond through the center of the ring to form the core of the knot. The 4th cysteine forms an interchain disulfide bond with the corresponding residue in the other subunit.

The TGF-β2 and OP-1 monomer subunits comprise three major structural elements and an N-terminal region. The structural elements are made up of regions of contiguous polypeptide chain that possess over 50% secondary structure of the following types: (1) loop, (2) α-helix and (3) β-sheet. Furthermore, in these regions the N-terminal and C-terminal strands are not more than 7 Å apart. The residues between the 1st and 2nd conserved cysteines form a structural region characterized by an anti-parallel β-sheet finger, referred to herein as the finger 1 region (F1). Similarly the residues between the 5th and 6th conserved cysteines also form an anti-parallel β-sheet finger, referred to herein as the finger 2 region (F2). A β-sheet finger is a single amino acid chain, comprising a β-strand that folds back on itself by means of a β-turn or some larger loop so that the entering and exiting strands form one or more anti-parallel β-sheet structures. The third major structural region, involving the residues between the 3rd and 4th conserved cysteines is characterized by a three turn α-helix referred to herein as the heel region (H). In the dimeric forms of both TGF-β2 and OP-1, the subunits are oriented such that the heel region of one subunit contacts the finger regions of the other subunit with the knot regions of the connected subunits forming the core of the molecule. The 4th cysteine forms a disulfide bridge with its counterpart on the second chain thereby equivalently linking the chains at the center of the palms, as further described herein below. The dimer thus formed is an ellipsoidal (cigar shaped) molecule when viewed from the top looking down the two-fold axis of symmetry between the subunits.

Whether naturally-occurring, or recombinantly or synthetically prepared, true morphogens within the TGF-β superfamily, such as osteogenic proteins, can induce recruitment and stimulation of progenitor cells, thereby inducing their differentiation, e.g., into chondrocytes and osteoblasts, and further inducing differentiation of intermediate cartilage, vascularization, bone formation, remodeling, and, finally, marrow differentiation. Numerous practitioners have demonstrated the ability of osteogenic proteins, when admixed with either naturally-sourced

matrix materials such as collagen or synthetically-prepared polymeric matrix materials, to induce true bone formation, including endochondral bone formation, under conditions where true replacement bone would not otherwise occur. For example, when combined with a matrix material, these osteogenic proteins induce formation of new bone in large segmental bone defects, spinal fusions, and fractures, to name but a few.

Bacterial and other prokaryotic expression systems are relied on in the art as preferred means for generating both native and biosynthetic or recombinant proteins. Prokaryotic systems such as *E. coli* are useful for producing commercial quantities of proteins, as well as for evaluating biological and chemical properties of naturally occurring or synthetically prepared mutants and analogs. Typically, an over-expressed eukaryotic protein aggregates into an insoluble intracellular precipitate ("inclusion body") in the prokaryote host cell. The aggregated protein is then collected from the inclusion bodies, solubilized using one or more standard denaturing agents, and then allowed, or induced, to refold into a functional state.

Optimal refolding requires proper formation of any disulfide bonds to form and stabilize a biologically active protein structure. However, not all naturally-occurring proteins when solubilized readily refold. Despite careful manipulation of the chemistries for refolding, the yields of optimally folded, stable and/or biologically active protein remain low. Many of the aforementioned proteins, including BMPs, fall into the category of poor refolder proteins. For example, while some members of the BMP protein family can be folded relatively efficiently *in vitro* as, for example, when produced in *E. coli* or other prokaryotic hosts, many others, including BMP-5, BMP-6, and BMP-7, are not. See, e.g., EP 0433225, US 5,399,677, 5,756,308, and 5,804,416.

A need remains for improved means for designing and successfully producing recombinant, chemosynthetic and/or biosynthetic members of the TGF-β superfamily of proteins, including morphogenic proteins.

Summary of the Invention

The present invention provides chimeric proteins, and DNAs encoding the same, of the TGF-β superfamily including morphogenic proteins. As used herein, the terms TGF-β superfamily "chimeric protein", "chimera", "chimeric polypeptide chain", "chimeric construct"

and "chimeric mutant" refer to any TGF-β superfamily member synthetic construct wherein the amino acid sequence (or corresponding nucleic acid encoding the same) of at least one defined region, domain or sub-domain, such as the finger 1, finger 2 or heel sub-domain, has been altered in whole or in part by exchanging it with an amino acid sequence from at least one other, different TGF-β superfamily protein, such that the resulting construct has an amino acid sequence which is non-naturally occurring and derived from at least two different protein sources. This structural alteration(s) is referred to herein as "domain-swapping." As contemplated herein, chimeric constructs also comprise recombinant and/or biosynthetic and/or chemosynthetic proteins. Additionally, chimeric proteins of the invention have altered refolding attributes relative to naturally occurring proteins. Chimeric proteins of the invention can have altered stability, solubility, surface binding, bioactivity and/or biospecificity attributes, and are particularly amendable to tissue-specific targeting due to their altered properties.

As set forth herein, it has been discovered that individual subdomains of the C-terminal active region of a first TGF-β superfamily member, such as a first BMP, can be exchanged with that of a second family member, such as a second BMP, to acquire a desired biological or biochemical attribute. In one embodiment, the invention provides synthetic forms of proteins that improve the refolding properties of the native "poor refolder" proteins. As used herein, a "poor refolder" protein means any protein that, when induced to refold under typically suitable refolding conditions, yields less than about 1% optimally refolded material (see below). Further, it has been found that the biological attribute(s) of a good refolder can be altered by exchanging the finger 1 and/or heel subdomains, without substantially affecting or compromising the folding properties of the parent protein. As exemplified herein, the finger 1 subdomain can act as a signature sequence conferring altered bioactivity and/or altered biospecificity.

As a result of these discoveries, means are now available for predicting and designing *de novo* TGF-β superfamily member chimerics having altered attributes, including altered folding capabilities, altered solubility, altered stability, altered isoelectric points, and/or altered biological activities and specificities, as desired. It is the customized reconfiguration and mixing-and-matching of attributes such as these which contributes to the novelty and utility of the present invention. The invention also provides means for easily and quickly evaluating biological and/or biochemical attributes of chimeric constructs, including mapping epitopes.

Thus, in one aspect, the present invention is directed to a TGF-β superfamily chimeric protein derived from at least two different members of said superfamily, said chimeric protein comprising a dimer wherein one monomer comprises a finger 1 subdomain, a finger 2 subdomain and a heel subdomain, said finger 2 subdomain being derived from a first member of said superfamily, said finger 1 or heel subdomain being derived from a second, different member of said superfamily, wherein said monomer further comprises a conserved C-terminal cysteine skeleton. In a preferred embodiment, the finger 1 or heel subdomain is derived from the BMP-7, OP-1. In other preferred embodiments, the finger 2 domain is derived from CDMP-2 or BMP-2. It is further contemplated that the chimeras of the instant invention can be homodimeric or heterodimeric.

As used herein, a "poor refolder" protein means any protein that, when induced to refold under typically suitable refolding conditions, yields less than about 1% optimally refolded material, as measured using a standard protocol. As contemplated herein, "typically suitable refolding conditions" are conditions under which proteins can be refolded to the extent required to confer functionality. One skilled in the art will recognize that at least Section I.B.1.(c) and Example 3 disclose non-limiting examples of such refolding conditions. Structural parameters relevant to the compositions and methods of the instant invention include one or more disulfide bridges properly distributed throughout the dimeric protein's structure and which require a reduction-oxidation ("redox") reaction step to yield a folded structure. Redox reactions typically occur at neutral pHs. Accordingly, as used herein, "suitable refolding conditions" include a redox reaction step at a substantially neutral pH, i.e., in the range of about pH 5.0-10.0, typically in the range of about pH 6.0-9.0, and preferably under physiologically compatible conditions. The skilled artisan will appreciate and recognize optimal conditions for success.

In one preferred embodiment, the invention provides recombinant, chemosynthetic or biosynthetic TGF-β superfamily member proteins having altered refolding properties under neutral or otherwise physiologically-compatible conditions. By way of example only, the recombinant, chemosynthetic or biosynthetic proteins of the invention have altered refolding properties at a pH in the range of about 5.0-10.0, preferably in the range of about 6.0-9.0, more preferably in the range of about 7.0-8.5, including in the range of about pH 7.5-8.5.

In another embodiment, the invention provides recombinant, chemosynthetic or biosynthetic TGF-β superfamily member proteins having altered solubility attributes under neutral or otherwise physiologically compatible conditions. In one embodiment, the proteins of the invention have altered solubility at a pH in the range of about 5.0-10.0, preferably in the range of about 6.0-9.0, more preferably in the range of about 6.0-8.5, including in the range of about pH 7.0-7.5. In another embodiment, the stability of a protein is altered by the modifications and manipulations disclosed herein. 'Altered" is intended to mean different from the native protein's attribute(s), and includes embodiments which can be more stable or less stable, and/or more soluble or less soluble, etc.

In still another embodiment, the invention provides TGF-β superfamily constructs competent to refold under physiologically-compatible conditions and having altered isoelectric points as compared with the native protein. In another embodiment, the invention provides TGF-β superfamily member proteins having altered refolding properties as compared with the native protein, and wherein the constructs also have, for example, altered receptor binding specificity, altered stability, altered solubility and/or biological activity.

In another embodiment, the invention provides TGF-β superfamily chimerics having little or no substantial increase in immunogenic effect in a mammal as compared with the parent sequence.

Modified proteins of the invention can be used in conjunction with a biocompatible matrix such as, but not limited to, collagen, hydroxyapatite, ceramics or carboxymethylcellulose, or other suitable matrix material. Such combinations are particularly useful in methods for regenerating bone, cartilage and/or other non-mineralized skeletal or connective tissues such as but not limited to ligament, tendon, muscle, articular cartilage, fibrocartilage, joint capsule, menisci, intervertebral discs, synovial membrane tissue, and fasica to name but a few. See e.g. U.S. Patent No. 5,496,552, 5,674,292, 5,840,325 and U.S.S.N. 08/253,398, soon-to-issue as U.S. _____, the disclosures of which are incorporated by reference herein; also incorporated by reference herein are co-pending U.S.S.N. 08/459,129 and 08/458,811 each filed on June 2, 1995. The instant invention contemplates that the binding and/or adherence properties to such matrix materials can be altered using the domain swapping techniques disclosed herein.

In another aspect, the invention provides recombinant or synthetic nucleic acid molecules, including DNA, RNA and PNA (peptide-nucleic acid) molecules, encoding any of the chimeric TGF-β superfamily member proteins of the invention.

In yet another aspect, the invention provides novel methods for creating biosynthetic, chemosynthetic or recombinant constructs having altered biological attributes. In one embodiment, individual sub-domains of the C-terminal active region can be exchanged between molecules to create a chimeric construct with a desired property. For example, the biological activity of a good refolder can be altered, without substantially affecting the refolding properties of the protein, by replacing the parent protein's finger 1 sub-domain and/or heel sub-domain with that of another having the desired activity.

In another embodiment, the invention provides a method for folding those native TGF-β superfamily proteins which are poor refolders, such as BMP homodimers and heterodimers, as well as for folding the chimeric proteins of the present invention under physiologically compatible and/or neutral pH conditions. In one preferred embodiment, the method comprises the steps of providing one or more solubilized substituted BMPs of the invention, exposing the solubilized BMP to a redox reaction in a suitable refolding buffer, and allowing the protein subunits to refold into homodimers and/or heterodimers, as desired. In another embodiment, the redox reaction system can utilize oxidized and reduced forms of glutathione, DTT, β-mercaptoethanol, β-mercaptomethanol, cystine and cystamine. In another embodiment the redox reaction system relies on air oxidation, preferably in the presence of a metal catalyst, such as copper. In still another embodiment, ratios of reductant to oxidant of about 1:10 to about 10:1, preferably in the range of about 1:2 to 2:1, can be used. In another preferred embodiment, the chimeric protein is solubilized in the presence of a detergent, including a non-ionic detergent, e.g. digitonin, N-octyl glucoside, or zwitterionic detergents, such as 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfate (CHAPS). In still another embodiment, the refolding reaction occurs in a pH range of about 5.0-10.0, preferably in the range of about 6.0-9.0, more preferably in the range of about 7.0-8.5. In still another embodiment, the refolding reaction occurs at a temperature within the range of about 0–32°C, preferably in the range of about 4-25°C. Where heterodimers are being created, optimal ratios for adding the two different subunits readily can be determined empirically and without undue experimentation.

In yet another embodiment, the invention provides a method for determining the epitope of an antibody that binds to one member of a TGF-β superfamily ("first member") but not to a different member of the TGF-β superfamily ("second member"). A chimeric protein is created that replaces all or part of the C-terminal subdomain of the first member with that of the second member. An antibody is introduced to this first member/second member chimeric protein under conditions that would allow the antibody to bind to its antigen – the wild type first member. If the antibody does bind to the chimeric protein, it indicates that the epitope of the antibody is not within the part of first member that was replaced by the second member. However, if the antibody does not bind the chimeric protein, that indicates that the epitope for the antibody is in that region of the first member that was replaced by the second member.

Certain of the compositions of the present invention preferably are manufactured in accordance with the principles disclosed herein by assembly of nucleotides and/or joining DNA restriction fragments to produce synthetic DNAs. The DNAs are transfected into an appropriate protein expression vehicle, the encoded protein expressed, folded, and purified. Particular constructs can be tested for agonist activity *in vitro*. The tertiary structure of the candidate constructs can be iteratively refined and modulated by site-directed or nucleotide sequence directed mutagenesis aided by the principles disclosed herein, computer-based protein structure modeling, and recently developed rational molecular design techniques to improve or modulate specific properties of a molecule of interest. Known phage display or other expression systems can be exploited to produce simultaneously a large number of candidate constructs. The pool of candidate constructs subsequently can be screened for altered and/or improved binding specificity using, for example, a chromatography column comprising surface immobilized receptors, salt gradient elution to select for, and to concentrate high binding candidates, and *in vitro* assays to determine whether or not particular isolated candidates agonize the activity of the template superfamily member(s). Identification of a useful construct is followed by production of cell lines expressing commercially useful quantities of the construct for laboratory use and ultimately for producing therapeutically useful molecules. It is contemplated also that preferred constructs, once identified and characterized by the protein and DNA methodologies described herein, can be produced by standard chemical synthesis methodologies.

In view of the fundamental know-how discovered and described herein, it is now possible

to design, make, test and use chimeric proteins, including modified morphogenic proteins, comprising dimeric amino acid sequences which, when optimally folded, assume a tertiary structure defining a finger 1 region, a finger 2 region, and a heel region which together are complementary to the ligand binding interactive surface of a TGF-β superfamily member receptor. The proteins, having acquired the ability to bind with a particular receptor, for example, can agonize the activity of the desired TGF-β superfamily member. In one subset of embodiments, the proteins can agonize initiation of cellular differentiation and tissue morphogenesis, e.g., initiate cell transformation leading to new tissue formation, such as bone formation. In another embodiment, the proteins comprise amino acid sequences sufficiently duplicative of the amino acid sequences of a preferred TGF-β superfamily member such that it can now preferentially bind the cognate receptor for that preferred member.

All of the protein constructs, and DNAs corresponding thereto, of the invention comprise regions of amino acid sequences defining the three regions required for utility, namely, finger 1, finger 2, and heel regions, maintained in their proper conformation and maintained in their relative positions and orientations in space. Sequences for the finger and heel regions can correspond to the respective finger and heel region sequences of any known TGF-β superfamily member identified herein. Alternatively, the finger and heel regions can be selected from the amino acid sequence of a new member of this superfamily discovered hereafter using the principles disclosed hereinbelow.

The finger and heel sequences also can be altered by amino acid substitution, for example by exploiting substitute amino acid residues selected in accordance with the principles disclosed in Smith et al. (1990) Proc. Natl. Acad. Sci. USA 87: 118-122, the disclosure of which is incorporated herein by reference. Smith et al. disclose an amino acid class hierarchy, similar to the amino acid hierarchy table set forth in Figure 8, which may be used to rationally substitute one amino acid for another while minimizing gross conformational distortions of the type which otherwise may inactivate the protein. In any event, it is contemplated that many synthetic finger 1, finger 2, and heel region sequences, having only 70% homology with natural regions, preferably 80%, and most preferably at least 90%, can be used to produce active constructs. It is contemplated also, as disclosed herein, that the size of the protein constructs can be reduced significantly by truncating the N-terminal region upstream of the first cysteine or the natural

finger and heel regions of the template TGF-β superfamily member.

The foregoing and other objects, features and advantages of the present invention will be made more apparent from the following detailed description of preferred embodiments of the invention.

Brief Description of the Drawings

Figure 1 lists the aligned C-terminal residues defining the finger 2 sub-domain for various known members of the BMP family, and TGF-β superfamily of proteins, starting with the first residue following the cysteine doublet.

Figure 2 is a schematic representation of a ribbon model of a BMP C-terminal active domain, showing the seven cysteine skeleton and the three distinct sub-domains: heel (10), finger 1 (12), and finger 2 (14), including the neck or base region (16) and the tip or loop region (18); the half-disulfide link at cysteine 4 crosslinks the monomer of Figure 2 to a second monomer to yield the dimeric form.

Figure 3 is a nucleotide sequence and the corresponding amino acid sequence of the OP-1 C-terminal seven cysteine active domain.

Figure 4(A) is a schematic representation of exemplary embodiments of chimeric protein constructs of the present invention.

Figure 4(B) is a schematic representation of exemplary embodiments of the present invention relating to attributes of refolding and ROS activity.

Figure 5 illustrates a correlation between refolding efficiency and charged amino acids in the TGF-β 7-cysteine domain.

Figures 6A, 6B, and 6C are sequence alignments using single letter amino acid code, arranged to indicate homologies of the finger 1, heel, and finger 2 regions, respectively, of some known members of the TGF-β superfamily. Shown are the respective amino acids comprising each region of human TGF-β1 through TGF-β5 (the TGF-β subgroup), the Vg/dpp subgroup consisting of dpp, Vg-1, Vgr-1, 60A (see copending U.S.S.N. 08/271,556), BMP-2A (also known in the literature as BMP-2), dorsalin, BMP-2B (also known in the literature as BMP-4), BMP-3, BMP-5, BMP-6, OP-1 (also known in the literature as BMP-7), OP-2 (see

PCT/US91/07635 and U.S. Patent No. 5,266,683) and OP-3 (U.S.S.N 07/971,091), the GDF subgroup consisting of GDF-1, GDF-3, and GDF-9, the Inhibin subgroup consisting of Inhibin α, Inhibin βA, and Inhibin βB. The dashes (-) indicate a peptide bond between adjacent amino acids. A consensus sequence pattern for each subgroup is shown at the bottom of each subgroup.

Figure 7 is a single letter code listing of amino acid sequences, identified in capital letter in standard single letter amino acid code, and in lower case letters to identify groups of amino acids useful in that location, wherein the lower case letters stand for the amino acids indicated in accordance with the pattern definition key table set forth in Figure 8. Figure 7 identifies preferred pattern sequences for constituting the finger 1, heel, and finger 2 regions of biosynthetic constructs of the invention. The dashes (-) indicate a peptide bond between adjacent amino acids.

Figure 8 is a pattern definition table prepared in accordance with the teaching of Smith and Smith (1990) Proc. Natl. Acad. Sci. USA 87:118-122.

Detailed Description of Preferred Embodiments

Before proceeding further with a detailed description of the currently preferred embodiments of the instant invention, an explanation of certain terms and phrases will be provided. Accordingly, it is understood that each of the terms or phrases set forth below is defined herein at least as follows:

As used herein, "acidic" or "negatively charged residues" are understood to include any amino acid residue, naturally-occurring or synthetic, that carries a negative charge on its R group under neutral pHs, including but not limited to physiologically compatible conditions. Examples include, without limitation, aspartic acid ("Asp") and glutamic acid ("Glu"). Similarly, basic or positively charged residues include any amino acid residue, naturally-occurring or synthetically created, that carries a positive charge on its R group under physiological conditions. Examples include, without limitation, arginine ("Arg"), lysine ("Lys") and histidine ("His"). As used herein, "hydrophilic" residues include both acidic and basic amino acid residues, as well as uncharged residues carrying amide groups on their R groups, including, without limitation, glutamine ("Gln") and asparagine ("Asn"), and polar residues carrying hydroxyl groups on their R groups, including, without limitation, serine ("Ser") and threonine ("Thr").

As used herein, "biosynthesis" or "biosynthetic" means occurring as a result of, or originating from ligation of naturally-or synthetically-derived fragments. For example, but not limited to, ligating peptide or nucleic acid fragments corresponding to one or more subdomains (or fragments thereof) disclosed herein. "Chemosynthesis" or "chemosynthetic" means occurring as a result of, or originating from, a chemical means of production. For example, but not limited to, synthesis of a peptide or nucleic acid sequence using a standard automated synthesizer from a commercially-available source. It is contemplated that both natural and non-natural amino acids can be used to obtain the desired attributes as taught herein. "Recombinant" production or technology means occurring as a result of, or originating from, a genetically engineered means of production. For example, but not limited to, expression of a genetically-engineered DNA sequence or gene encoding a chimeric protein (or fragment thereof) of the present invention.. Also included within the meaning of the foregoing are the teachings set forth below in at least Sections I.B.1. (a) and (b); Section II; and at least Examples 1, 2 and 9 (Section III). "Synthetic" means occurring or originating non-naturally, i.e., not naturally occurring.

As used herein, "corresponding residue position" refers to a residue position in a protein sequence that corresponds to a given position in a reference amino acid sequence, when the two sequences are aligned. As will be appreciated by those skilled in the art and as illustrated in Fig.1, the sequences of the BMP family members are highly conserved in the C-terminal active domain, and particularly in the finger 2 sub-domain. Amino acid sequence alignment methods and programs are well developed in the art. See, e.g., the method of Needleman, et al. (1970) *J. Mol. Biol. 48*:443-453, implemented conveniently by computer programs such as the Align program (DNAstar, Inc.). Internal gaps and amino acid insertions in the second sequence are ignored for purposes of calculating the alignment. For ease of description, hOP-1 (human OP-1, also referred to in the art as "BMP-7") is provided below as a representative TGF-β superfamily member. It will be appreciated however, that OP-1 is merely representative of the TGF-β subclass of true tissue morphogens competent to induce tissue morphogenesis.

"Osteogenic protein", or "bone morphogenic protein," means a TGF-β superfamily protein which can induce the full cascade of morphogenic events culminating in skeletal tissue formation, including but not limited to cartilage and/or endochondral bone formation. Osteogenic proteins useful herein include any known naturally-occurring native proteins

including allelic, phylogenetic counterpart and other variants thereof, whether naturally-occurring or biosynthetically produced (*e.g.*, including "muteins" or "mutant proteins"), as well as new, osteogenically active members of the general morphogenic family of proteins. As described herein, this class of proteins is generally typified by human osteogenic protein 1 (hOP-1). Other osteogenic proteins useful in the practice of the invention include osteogenically active forms of proteins included within the list of: OP-1, OP-2, OP-3, BMP-2, BMP-3, BMP-4, BMP-5, BMP-6, BMP-9, DPP, Vg-1, Vgr, 60A protein, CDMP-1, CDMP-2, CDMP-3, GDF-1, GDF-3, GDF-5, 6, 7, MP-52, BMP-10, BMP-11, BMP-12, BMP-13, BMP-15, UNIVIN, NODAL, SCREW, ADMP or NEURAL, including amino acid sequence variants thereof, and/or heterodimers thereof. In one currently preferred embodiment, osteogenic protein useful in the practice of the invention includes any one of: OP-1, BMP-2, BMP-3, BMP-4, BMP-5, BMP-6, BMP-12, BMP-13, GDF-5, GDF-6, GDF-7, CDMP-1, CDMP-2, CDMP-3, MP-52 and amino acid sequence variants and homologs thereof, including species homologs thereof. In still another preferred embodiment, useful osteogenically active proteins have polypeptide chains with amino acid sequences comprising a sequence encoded by a nucleic acid that hybridizes, under low, medium or high stringency hybridization conditions, to DNA or RNA encoding reference osteogenic sequences, *e.g.*, C-terminal sequences defining the conserved seven cysteine domains of OP-1, OP-2, BMP-2, BMP-4, BMP-5, BMP-6, 60A, GDF-5, GDF-6, GDF-7 and the like. As used herein, medium stringent hybridization conditions are defined as hybridization according to known techniques in 40% formamide, 5 X SSPE, 5 X Denhardt's Solution, and 0.1% SDS at 37°C overnight, and washing in 0.1 X SSPE, 0.1% SDS at 50°C. Standard stringency conditions are well characterized in commercially available, standard molecular cloning texts. See, for example, Molecular Cloning A Laboratory Manual, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); DNA Cloning, Volumes I and II (D.N. Glover ed., 1985); Oligonucleotide Synthesis (M.J. Gait ed., 1984): Nucleic Acid Hybridization (B. D. Hames & S.J. Higgins eds. 1984); and B. Perbal, A Practical Guide To Molecular Cloning (1984); the disclosures of the foregoing are incorporated by reference herein. See also, U.S. Patent Nos. 5,750,651 and 5,863,758, the disclosures of which are incorporated by reference herein.

Other members of the TGF-ß superfamily of related proteins having utility in the practice of the instant invention include native poor refolder proteins among the list: TGF-β1, TGF-β2, TGF-β3, TGF-β4 and TGF-β5, various inhibins, activins, BMP-11 and MIS, to name a few. Fig. 1 lists the C-terminal residues defining the finger 2 subdomain, along with the flanking cysteines, of various known members of the TGF-ß superfamily. Any one of the proteins on the list that is a poor refolder can be improved by the methods of the invention, as can other known or discoverable family members. As further described herein, the biologically active osteogenic proteins suitable for use with the present invention can be identified by means of routine experimentation using the bioassay described by Reddi and Sampath as recognized in the art. A detailed description of useful proteins follows. Equivalents can be identified by the artisan using no more than routine experimentation and ordinary skill.

"Morphogens" or "morphogenic proteins" as contemplated herein includes members of the TGF-β superfamily which have been recognized to be morphogenic, i.e., capable of inducing the developmental cascade of tissue morphogenesis in a mature mammal (See PCT Application No. US 92/01968). In particular, these morphogens are capable of inducing the proliferation of uncommitted progenitor cells, and inducing the differentiation of these stimulated progenitor cells in a tissue-specific manner under appropriate environmental conditions. In addition, the morphogens are capable of supporting the growth and maintenance of these differentiated cells. These morphogenic activities allow the proteins to initiate and maintain the developmental cascade of tissue morphogenesis in an appropriate, morphogenically permissive environment, stimulating stem cells to proliferate and differentiate in a tissue-specific manner, and inducing the progression of events that culminate in new tissue formation. These morphogenic activities also allow the proteins to induce the "redifferentiation" of cells previously stimulated to stray from their differentiation path. Under appropriate environmental conditions it is anticipated that these morphogens also may stimulate the "redifferentiation" of committed cells. To guide the skilled artisan, described herein are numerous means for testing morphogenic proteins in a variety of tissues and for a variety of attributes typical of morphogenic proteins. It will be understood that these teachings can be used to assess morphogenic attributes of native proteins as well as chimeric proteins of the present invention. See, for example, Sections III, Examples 11A-11G for teachings related to morphogenesis of brain, liver, to name but a few.

Useful native or parent proteins of the present invention also include those sharing at least 70% amino acid sequence homology within the C-terminal seven-cysteine domain of human OP-1. To determine the percent homology of a candidate amino acid sequence to the conserved seven-cysteine domain, the candidate sequence and the seven cysteine domain are aligned. The first step for performing an alignment is to use an alignment tool, such as the dynamic programming algorithm described in Needleman *et al.*, J. Mol. Biol. *48*: 443 (1970), the teachings of which are incorporated by reference herein and the Align Program, a commercial software package produced by DNAstar, Inc. After the initial alignment is made, it is then refined by comparison to a multiple sequence alignment of a family of related proteins. Once the alignment between the candidate sequence and the seven-cysteine domain is made and refined, a percent homology score is calculated. The individual amino acids of each sequence are compared sequentially according to their similarity to each other. Similarity factors include similar size, shape and electrical charge. One particularly preferred method of determining amino acid similarities is the PAM250 matrix described in Dayhoff *et al.*, 5 Atlas of Protein Sequence and Structure 345-352 (1978 & Supp.), incorporated by reference herein. A similarity score is first calculated as the sum of the aligned pairwise amino acid similarity scores. Insertions and deletions are ignored for the purposes of percent homology and identity. Accordingly, gap penalties are not used in this calculation. The raw score is then normalized by dividing it by the geometric mean of the scores of the candidate compound and the seven cysteine domain. The geometric mean is the square root of the product of these scores. The normalized raw score is the percent homology.

As used herein, "conservative substitutions" are residues that are physically or functionally similar to the corresponding reference residues, *e.g.*, that have similar size, shape, electric charge, chemical properties including the ability to form covalent or hydrogen bonds, or the like. Particularly preferred conservative substitutions are those fulfilling the criteria defined for an accepted point mutation in Dayhoff *et al.* (1978), 5 Atlas of Protein Sequence and Structure, Suppl. 3, ch. 22 (pp. 354-352), Natl. Biomed. Res. Found., Washington, D.C. 20007, incorporated by reference herein. Examples of conservative substitutions include the substitution of one amino acid for another with similar characteristics, *e.g.*, substitutions within the following groups are well-known: (a) glycine, alanine; (b) valine, isoleucine, leucine; (c) aspartic acid,

glutamic acid; (d) asparagine, glutamine; (e) serine, threonine; (f) lysine, arginine, histidine; and (g) phenylalanine, tyrosine. The term "conservative variant" or "conservative variation" also includes the use of a substituted amino acid in place of an unsubstituted parent amino acid in a given polypeptide chain, provided that antibodies having binding specificity for the resulting substituted polypeptide chain also have binding specificity (i.e., "crossreact" or "immunoreact" with) for the unsubstituted or parent polypeptide chain.

As used herein, a "conserved residue position" refers to a location in a reference amino acid sequence occupied by the same amino acid or a conservative variant thereof in at least one other member sequence. For example, in Fig. 1, comparing BMP-2, BMP-4, BMP-5, and BMP-6 with OP-1 as the reference sequence, positions 2(P), 3(T), 5(L), 7(A), 8(I), and 9(S), etc. are conserved positions, and residues 4(K,E) and 6(N,S), etc. are non-conserved positions.

As used herein, the "base" or "neck" region of the finger 2 sub-domain is defined by residues 1-10 and 22-31, as exemplified by OP-1, and counting from the first residue following the cysteine doublet in the C-terminal active domain. (See Fig. 1). As is readily apparent from a sequence alignment of other TGF-β superfamily protein members with OP-1, the corresponding base or neck region for a longer protein, such as BMP-9 or Dorsalin, is defined by residues 1-10 and 23-32; for a shorter protein, such as NODAL, the corresponding region is defined by residues 1-10 and 22-30 (See Fig. 1). In SEQ ID NO: 39, (human OP-1), the residues corresponding to the base or neck region of the finger 2 subdomain are residues 397-406 (corresponding to residues 1-10 in Fig. 1) and residues 418-427 (corresponding to residues 22-31 in Fig. 1).

As used herein, "amino acid sequence homology" includes both amino acid sequence identity and similarity. Homologous sequences share identical and/or similar amino acid residues, where similar residues are conservative substitutions for, or "allowed point mutations" of, corresponding amino acid residues in an aligned reference sequence.

As used herein, the terms "chimeric protein", "chimera", "chimeric polypeptide chain", "chimeric construct" and "chimeric mutant" refer to any TGF-β superfamily member synthetic construct wherein the amino acid sequence of at least one defined region, domain or sub-domain, such as a first finger 1, finger 2 or heel sub-domain, has been replaced in whole or in part (e.g., at least about 3, 5, 10, or 15 consecutive amino acid residues) with a second amino acid sequence

from at least one other, different TGF-β superfamily member protein, such that the resulting construct has an amino acid sequence recognizable as non-naturally occurring and derived from the different protein sources. Certain preferred chimerics contain a sub-domain from a third different TGF-β superfamily protein or contain two different subdomains from the second protein. It is understood that a hallmark of preferred embodiments is a refolded structure in which the conserved TGF-β superfamily di-sulfide bridges are retained; see Section I.A. below for a discussion of the conserved 1st, 2nd, 3rd, 5th, 6th, and 7th cysteine residues *versus* the semi-conserved 4th residue.

As used herein, useful expression host cells include prokaryotes and eukaryotes, including any host cell capable of making an inclusion body. Particularly useful host cells include, without limitation, bacterial hosts such as *E. coli*, as well as *B. subtilis* and *Pseudomonas.* Other useful hosts include lower eukaryotes, such as yeast (*Saccharomyces cereviceae*) and higher eukaryotes, including insect cells such as *Drosophila,* mammalian cells, such as CHO and the like.

According to the present invention, the attributes of native BMPs or other members of the TGF-β superfamily of proteins, including heterodimers and homodimers thereof, are altered by domain-swapping, for example, by exchanging defined regions of the F1 subdomain between proteins to alter one or more biological properties of a BMP or TGF-β superfamily member. As a result of this discovery, it is possible to design, TGF-β superfamily proteins that (1) are expressed recombinantly in prokaryotic or eukaryotic cells synthesized using polypeptide or nucleotide synthesizers; (2) have altered folding attributes; (3) have altered solubility under neutral pHs, including but not limited to physiologically compatible conditions; (4) have altered isoelectric points; (5) have altered stability; (6) have an altered tissue or receptor specificity; (7) have a re-designed, altered biological activity; and/or (8) have altered binding or adherence properties to solid surfaces, such as but not limited to, biocompatible matrices or metals. Thus, the present invention can provide mechanisms for designing quick-release, slow-release and/or timed-release formulations containing a preferred chimeric protein. Other advantages and features will be evident from the teachings below. Moreover, making use of the discoveries disclosed herein, chimeric proteins having altered surface-binding/surface-adherent properties

can be designed and selected. Surfaces of particular significance include, but are not limited to, solid surfaces which can be naturally-occurring such as bone; or porous particulate surfaces such as collagen or other biocompatible matrices; or the fabricated surfaces of prosthetic implants, including metals. As contemplated herein, virtually any surface can be assayed for differential binding of chimeric constructs. Thus, the present invention embraces a diversity of functional molecules having alterations in their surface-binding/surface-adherent properties, thereby rendering such constructs useful for altered *in vivo* applications, including slow-release, fast-release and/or timed-release formulations.

The skilled artisan will appreciate that mixing-and-matching any one or more the above-recited attributes provides specific opportunities to manipulate the uses of customized chimeric proteins (and DNAs encoding the same). For example, the attribute of altered stability can be exploited to manipulate the turnover of a chimeric protein *in vivo*. Moreover, in the case of chimeric proteins also having attributes such as altered re-folding and/or function, there is likely an interconnection between folding, function and stability. See, for example, Lipscomb et al., 7 Protein Sci. 765-73 (1998); and Nikolova et al., 95 Proc. Natl. Acad. Sci. USA 14675-80 (1998). For purposes of the present invention, stability alterations can be routinely monitored using well-known techniques of circular dichroism other indices of stability as a function of denaturant concentration or temperature. One can also use routine scanning calorimetry. Similarly, there is likely an interconnection between any of the foregoing attributes and the attribute of solubility. In the case of solubility, it is possible to manipulate this attribute so that a chimeric protein is either more or less soluble under physiologically-compatible and it consequently diffuses readily or remains localized, respectively, when administered *in vivo*.

Provided below are detailed descriptions of suitable recombinant, chemosynthetic or biosynthetic proteins and DNAs, as well as methods useful in the practice of the invention, and methods for using and testing proteins constructs; also, provided below are numerous, nonlimiting examples which 1) illustrate the utility of the recombinant, chemosynthetic or biosynthetic proteins, DNAs and methods described herein; and 2) provide assays with which to test and use these protein and DNA constructs.

I. PROTEIN CONSIDERATIONS

A. Biochemical, Structural and Functional Properties of Exemplary Members of the TGF-β Superfamily.

Each of the subunits in either TGF β2 or OP-1 have a characteristic folding pattern, that involves six of the seven C-terminal cysteine residues. Briefly, four of the cysteine residues in each subunit form two disulfide bonds which together create an eight residue ring, while two additional cysteine residues form a disulfide bond that passes through the ring to form a knot-like structure. With a numbering scheme beginning with the most N-terminal cysteine of the 7 conserved cysteine residues assigned number 1, the 2nd and 6th cysteine residues are disulfide bonded to close one side of the eight residue ring while the 3rd and 7th cysteine residues are disulfide bonded to close the other side of the ring. The 1st and 5th conserved cysteine residues are disulfide bonded through the center of the ring to form the core of the knot. Amino acid sequence alignment patterns suggest this structural motif is conserved between members of the TGF-β superfamily. The 4th cysteine is semi-conserved and when present typically forms an interchain disulfide bond (ICDB) with the corresponding cysteine residue in the other subunit.

The structure of each subunit in TGF-β2 and OP-1 comprise three major tertiary structural elements and an N-terminal region. The structural elements are made up of regions of contiguous polypeptide chain that possess over 50% secondary structure of the following types: (1) loop, (2) α-helix and (3) β-sheet. Another defining criterion for each structural region is that the entering (N-terminal) and exiting (C-terminal) peptide strands are fairly close together, being about 7 Å apart.

The amino acid sequence between the 1st and 2nd conserved cysteines, forms a structural region characterized by an anti-parallel β-sheet finger referred to herein as the finger 1 region. Similarly the residues between the 5th and 6th conserved cysteines, also form an anti-parallel β-sheet finger, referred to herein as the finger 2 region. A β-sheet finger is a single amino acid chain, comprising a β-strand that folds back on itself by means of a β-turn or some larger loop so that the polypeptide chain entering and exiting the region form one or more anti-parallel β-sheet structures. The third major structural region, involving the residues between the 3rd and 5th conserved cysteines, is characterized by a three turn α-helix, referred to herein as the heel region.

The organization of the monomer structure is similar to that of a left hand where the knot region is located at the position equivalent to the palm, the finger 1 region is equivalent to the index and middle fingers, the α-helix, or heel region, is equivalent to the heel of the hand, and the finger 2 region is equivalent to the ring and small fingers. The N-terminal region, whose sequence is not conserved across the TGF-β superfamily, is predicted to be located at a position roughly equivalent to the thumb.

In a conformationally active TGF-β2 dimer complex, the two monomer subunits in the dimer complex are oriented with two-fold rotational summetry such that the heel region of one subunit contacts the finger regions of the other subunit with the knot regions of the connected subunits forming the core of the molecule. The 4th cysteine forms an interchain disulfide bond with its counterpart on the second chain thereby equivalently linking the chains at the center of the palms. The dimer thus formed is an ellipsoidal (cigar shaped) molecule when viewed from the top looking down the two-fold axis of symmetry between the subunits. Viewed from the side, the molecule resembles a bent "cigar" since the two subunits are oriented at a slight angle relative to each other.

The heel region from one subunit, and the finger 1 and finger 2 regions, respectively from the other subunit, interact with one another. These three elements co-operate with one other to define a structure interactive with, and complimentary to the ligand binding interactive surface of the cognate receptor.

Selection of Finger and Heel Regions

It is contemplated that the amino acid sequences defining the finger and heel regions can correspond to the respective finger and heel region sequences of any known member of the TGF-β superfamily. Furthermore, it is contemplated also that the size of the constructs of the invention can be reduced significantly by truncating the natural finger and heel regions of the template TGF-β superfamily member.

As mentioned above, amino acid sequences defining finger and heel regions can be from the respective finger and heel region sequences of any known TGF-β superfamily member,

identified herein, or from amino acid sequences of a new superfamily member discovered hereafter, can be used in the present invention.

Fig. 6 summarizes the amino acid sequences of currently identified TGF-β superfamily members aligned into finger 1 (Fig. 6A), heel (Fig. 6B) and finger 2 (Fig. 6C) regions. The sequences were aligned by a computer algorithm which in order to optimally align the sequences inserted gaps into regions of amino acid sequence known to define loop structures rather than regions of amino acid sequence known to have conserved amino acid sequence or secondary structure. For example, if possible, no gaps were introduced into amino acid sequences of finger 1 and finger 2 regions defined by β sheet or heel regions defined by α helix. The dashes (-) indicate a peptide bond between adjacent amino acids. A consensus sequence pattern for each subgroup is shown at the bottom of each subgroup.

After the amino acid sequences of each of the TGF-β superfamily members were aligned, the aligned sequences were used to produce amino acid sequence alignment patterns which identify amino acid residues that may be substituted by another amino acid or group of amino acids without altering the overall tertiary structure of the resulting construct. The amino acids or groups of amino acids that may be useful at a particular position in the finger and heel regions were identified by a computer algorithm implementing the amino acid hierarchy pattern structure shown in Fig. 8.

Briefly, the algorithm performs four levels of analysis. In level I, the algorithm determines whether a particular amino acid residue occurs with a frequency greater than 75% at a specific position within the amino acid sequence. For example, if a glycine residue occurs 8 out of 10 times at a particular position in an amino acid sequence, then a glycine is designated at that position. If the position to be tested consists of all gaps then a gap character (-) is assigned to the position, otherwise, if at least one gap exists then a "z" (standing for any residue or a gap) is assigned to the position. If, no amino acid occurs in 75% of the candidate sequences at a particular position the algorithm implements the Level II analysis.

Level II defines pattern sets a, b, d, l, k, o, n, i, and h, wherein l, k, and o share a common amino acid residue. The algorithm then determines whether 75% or more of the amino acid residues at a particular position in the amino acid sequence satisfy one of the aforementioned patterns. If so, then the pattern is assigned to that position. It is possible, however, that both

patterns l and k may be simultaneously satisfied because they share the same amino acid, specifically aspartic acid. If simultaneous assignment of l and k occurs then pattern m (Level III) is assigned to that position. Likewise, it is possible that both patterns k and o may be simultaneously assigned because they share the same amino acid, specifically glutamic acid. If simultaneous assignment of k and o occurs, then pattern q (Level III) is assigned to that position. If neither a Level II pattern nor the Level III patterns, m and q, satisfy a particular position in the amino acid sequence then the algorithm implements a Level III analysis.

Level III defines pattern sets c, e, m, q, p, and j, wherein m, q, and p share a common amino residue. Pattern q, however, is not tested in the Level III analysis. It is possible that both patterns m and p may be simultaneously satisfied because they share the same amino acid, specifically, glutamic acid. If simultaneous assignment of m and p occurs then pattern r (Level IV) is assigned to that position. If 75 % of the amino acids at a pre-selected position in the aligned amino acid sequences satisfy a Level III pattern, then the Level III pattern is assigned to that position. If a Level III pattern cannot be assigned to that position then the algorithm implements a Level IV analysis.

Level IV comprises two non-overlapping patterns f and r. If 75% of the amino acids at a particular position in the amino acid sequence satisfy a Level IV pattern then the pattern is assigned to the position. If no Level IV pattern is assigned the algorithm assigns an X representing any amino acid (Level V) to that position.

In Fig. 8, Level I lists in upper case letters in single amino acid code the 20 naturally occurring amino acids. Levels II-V define, in lower case letters, groups of amino acids based upon the amino acid hierarchy set forth in Smith et al., *supra*. The amino acid sequences set forth in Fig. 6 were aligned using the aforementioned computer algorithms.

It is contemplated that if the artisan wishes to produce a construct based upon currently identified members of the TGF-β superfamily, then the artisan can use the amino acid sequences shown in Fig. 6 to provide the finger 1, finger 2 and heel regions useful in the production of the chimerics of the invention. In the case of members of the TGF-β superfamily discovered hereafter, the amino acid sequence of the new member can be aligned, either manually or by means of a computer algorithm, with the sequences set forth in Fig. 6 to define heel and finger regions useful in the practice of the invention.

Table 1 below summarizes publications which describe the amino acid sequences of each TGF-β superfamily member that were used to produce the sequence alignment patterns set forth in Figs. 6-8.

Table 1.

TGF-β Superfamily Member	SEQ. ID. NO:	Publication
TGF-β1	40	Derynck et al. (1987) Nucl. Acids. Res. 15:3187
TGF-β2	41	Burt et al. (1991) DNA Cell Biol. 10:723-734
TGF-β3	42	Ten Dijke et al. (1988) Proc. Natl. Acad. Sci. USA 85:4715-4719; Derynck et al. (1988) EMBO J. 7:3737-3743.
TGF-β4	43	Burt et al. (1992) Mol. Endcrinol. 6:989-922.
TGF-β5	44	Kondaiah et al. (1990) J. Biol. Chem 265:1089-1093
dpp	45	Padgett et al. (1987) Nature 325:81-84; Paganiban et al. (1990) Mol. Cell Biol. 10:2669-2677.
vg-1	46	Weeks et al. (1987) Cell 51:861-867
vgr-1	47	Lyons et al. (1989) Proc. Natl. Acad. Sci USA 86:4554-4558
60A	48	Wharton et al. (1991) Proc. Natl. Acad. Sci. USA 88:9214-9218; Doctor et al. (1992) Dev. Biol. 151:491-505
BMP-2A	49	Wozney et al. (1988) Science 242: 1528-1534
BMP-3	50	Wozney et al. (1988) Science 242: 1528-1534
BMP-4	51	Wozney et al. (1988) Science 242: 1528-1534
BMP-5	52	Celeste et al. (1990) Proc. Natl. Acad.Sci. USA 87: 9843-9847
BMP-6	53	Celeste et al. (1990) Proc. Natl. Acad.Sci. USA 87: 9843-9847
Dorsalin	54	Basler et al. (1993) Cell 73:687-702

OP-1	55	Ozkaynak et al. (1990) EMBO J. 9:2085-2093; Celeste et al. (1990) Proc. Natl. Acad.Sci. USA 87: 9843-9847
OP-2	56	Ozkaynak et al. (1992) J. Biol. Chem.267: 25220-25227
OP-3	57	Ozkaynak et al. PCT/WO94/10203 Seq. I.D. No. 1.
GDF-1	58	Lee (1990) Mol. Endocrinol. 4: 1034-1040
GDF-3	59	McPherron et al. (1993) J. Biol. Chem. 268:3444-3449
GDF-9	60	McPherron et al. (1993) J. Biol. Chem. 268:3444-3449
Inhibin α	61	Mayo et al. (1986) Proc. Natl. Acad. Sci. USA 83:5849-5853; Stewart et al. (1986) FEBS Lett 206:329-334; Mason et al. (1986) Biochem. Biophys. Res. Commun. 135: 957-964
Inhibin βA	62	Forage et al. (1986) Proc. Natl. Acad. Sci. USA 83:3091-3095; Chertov et al. (1990) Biomed. Sci. 1:499-506
Inhibin βB	63	Mason et al. (1986) Biochem. Biophys. Res. Commun. 135:957-964

In particular, it is contemplated that amino acid sequences defining finger 1 regions useful in the practice of the instant invention correspond to the amino acid sequence defining a finger 1 region for any TGF-β superfamily member identified herein. The finger 1 subdomain can confer at least biological and/or functional attribute(s) which are characteristic of the native protein. Useful intact finger 1 regions include, but are not limited to

TGF-β1	SEQ ID NO: 40, residues 2 through 29,
TGF-β2	SEQ ID NO: 41, residues 2 through 29,
TGF-β3	SEQ ID NO: 42, residues 2 through 29,
TGF-β4	SEQ ID NO: 43, residues 2 through 29,
TGF-β5	SEQ ID NO: 44, residues 2 through 29,
dpp	SEQ ID NO: 45, residues 2 through 29,
Vg-1	SEQ ID NO: 46, residues 2 through 29,
Vgr-1	SEQ ID NO: 47, residues 2 through 29,
60A	SEQ ID NO: 48, residues 2 through 29,
BMP-2A	SEQ ID NO: 49, residues 2 through 29,

BMP-3	SEQ ID NO: 50, residues 2 through 29,
BMP-4	SEQ ID NO: 51, residues 2 through 29,
BMP-5	SEQ ID NO: 52, residues 2 through 29,
BMP-6	SEQ ID NO: 53, residues 2 through 29,
Dorsalin	SEQ ID NO: 54, residues 2 through 29,
OP-1	SEQ ID NO: 55, residues 2 through 29,
OP-2	SEQ ID NO: 56, residues 2 through 29,
OP-3	SEQ ID NO: 57, residues 2 through 29,
GDF-1	SEQ ID NO: 58, residues 2 through 29,
GDF-3	SEQ ID NO: 59, residues 2 through 29,
GDF-9	SEQ ID NO: 60, residues 2 through 29,
Inhibin α	SEQ ID NO: 61, residues 2 through 29,
Inhibin βA	SEQ ID NO: 62, residues 2 through 29,
Inhibin βB	SEQ ID NO: 63, residues 2 through 29,
CDMP-1/GDF-5	SEQ ID NO: 83, residues 2 through 29,
CDMP-2/GDF-6	SEQ ID NO: 84, residues 2 through 29,
GDF-6 (murine)	SEQ ID NO: 85, residues 2 through 29,
CDMP-2 (bovine)	SEQ ID NO: 86, residues 2 through 29, and
GDF-7 (murine)	SEQ ID NO: 87, residues 2 through 29.

The invention further contemplates the use of corresponding finger 1 subdomain sequences from the well-known proteins BMP-12 and BMP-13 (as disclosed in U.S. Patent No. 5,658,882, the entire disclosure of which is incorporated by reference herein).

It is contemplated also that amino acid sequences defining heel regions useful in the practice of the instant invention correspond to the amino acid sequence defining an intact heel region for any TGF-β superfamily member identified herein. The heel region can at least influence attributes of the native protein, including functional and/or folding attributes. Useful intact heel regions may include, but are not limited to

TGF-β1	SEQ ID NO: 40, residues 35 through 62,
TGF-β2	SEQ ID NO: 41, residues 35 through 62,
TGF-β3	SEQ ID NO: 42, residues 35 through 62,
TGF-β4	SEQ ID NO: 43, residues 35 through 62,
TGF-β5	SEQ ID NO: 44, residues 35 through 62,
dpp	SEQ ID NO: 45, residues 35 through 65,

Vg-1	SEQ ID NO: 46, residues 35 through 65,
Vgr-1	SEQ ID NO: 47, residues 35 through 65,
60A	SEQ ID NO: 48, residues 35 through 65,
BMP-2A	SEQ ID NO: 49, residues 35 through 64,
BMP-3	SEQ ID NO: 50, residues 35 through 66,
BMP-4	SEQ ID NO: 51, residues 35 through 64,
BMP-5	SEQ ID NO: 52, residues 35 through 65,
BMP-6	SEQ ID NO: 53, residues 35 through 65,
Dorsalin	SEQ ID NO: 54, residues 35 through 65,
OP-1	SEQ ID NO: 55, residues 35 through 65,
OP-2	SEQ ID NO: 56, residues 35 through 65,
OP-3	SEQ ID NO: 57, residues 35 through 65,
GDF-1	SEQ ID NO: 58, residues 35 through 70,
GDF-3	SEQ ID NO: 59, residues 35 through 64,
GDF-9	SEQ ID NO: 60, residues 35 through 65,
Inhibin α	SEQ ID NO: 61, residues 35 through 65,
Inhibin βA	SEQ ID NO: 62, residues 35 through 69,
Inhibin βB	SEQ ID NO: 63, residues 35 through 68,
CDMP-1/GDF-5	SEQ ID NO: 83, residues 35 through 65,
CDMP-2/GDF-6	SEQ ID NO: 84, residues 35 through 65,
GDF-6 (murine)	SEQ ID NO: 85, residues 35 through 65,
CDMP-2 (bovine)	SEQ ID NO: 86, residues 35 through 65, and
GDF-7 (murine)	SEQ ID NO: 87, residues 35 through 65.

The invention further contemplates the use of corresponding heel region sequences from the well-known proteins BMP-12 and BMP-13 (as disclosed in U.S. Patent No. 5,658,882, the entire disclosure of which is incorporated by reference herein).

It is contemplated also that amino acid sequences defining finger 2 regions useful in the practice of the instant invention correspond to the amino acid sequence defining an intact finger 2 region for any TGF-β superfamily member identified herein. The finger 2 subdomain can confer at least folding attribute(s) which are characteristic of the native protein. Useful intact finger 2 regions may include, but are not limited to

TGF-β1	SEQ ID NO: 40, residues 65 through 94,
TGF-β2	SEQ ID NO: 41, residues 65 through 94,

TGF-β3	SEQ ID NO: 42, residues 65 through 94,
TGF-β4	SEQ ID NO: 43, residues 65 through 94,
TGF-β5	SEQ ID NO: 44, residues 65 through 94,
dpp	SEQ ID NO: 45, residues 68 through 98,
Vg-1	SEQ ID NO: 46, residues 68 through 98,
Vgr-1	SEQ ID NO: 47, residues 68 through 98,
60A	SEQ ID NO: 48, residues 68 through 98,
BMP-2A	SEQ ID NO: 49, residues 67 through 97,
BMP-3	SEQ ID NO: 50, residues 69 through 99,
BMP-4	SEQ ID NO: 51, residues 67 through 97,
BMP-5	SEQ ID NO: 52, residues 68 through 98,
BMP-6	SEQ ID NO: 53, residues 68 through 98,
Dorsalin	SEQ ID NO: 54, residues 68 through 99,
OP-1	SEQ ID NO: 55, residues 68 through 98,
OP-2	SEQ ID NO: 56, residues 68 through 98,
OP-3	SEQ ID NO: 57, residues 68 through 98,
GDF-1	SEQ ID NO: 58, residues 73 through 103,
GDF-3	SEQ ID NO: 59, residues 67 through 97,
GDF-9	SEQ ID NO: 60, residues 68 through 98,
Inhibin α	SEQ ID NO: 61, residues 68 through 101,
Inhibin βA	SEQ ID NO: 62, residues 72 through 102,
Inhibin βB	SEQ ID NO: 63, residues 71 through 101,
CDMP-1/GDF-5	SEQ ID NO: 83, residues 68 through 98,
CDMP-2/GDF-6	SEQ ID NO: 84, residues 68 through 98,
GDF-6 (murine)	SEQ ID NO: 85, residues 68 through 98,
CDMP-2 (bovine)	SEQ ID NO: 86, residues 68 through 98, and
GDF-7 (murine)	SEQ ID NO: 87, residues 68 through 98.

The invention further contemplates the use of corresponding finger 2 subdomain sequences from the well-known proteins BMP-12 and BMP-13 (as disclosed in U.S. Patent No. 5,658,882, the entire disclosure of which is incorporated by reference herein).

In addition, it is contemplated that the amino acid sequences of the respective finger and heel regions can be altered by amino acid substitution, for example by exploiting substitute residues as disclosed herein or selected in accordance with the principles disclosed in Smith et al. (1990), *supra*. Briefly, Smith et al. disclose an amino acid class hierarchy similar to the one

summarized in Fig. 8, which can be used to rationally substitute one amino acid for another while minimizing gross conformational distortions of the type which could compromise protein function. In any event, it is contemplated that many synthetic first finger, second finger, and heel region sequences, having only 70% homology with natural regions, preferably 80%, and most preferably at least 90%, can be used to produce the constructs of the present invention.

Amino acid sequence patterns showing amino acids preferred at each location in the finger and heel regions, deduced in accordance with the principles described in Smith et al. (1990) supra, also are show in Figs. 6-7, and are referred to as the: TGF-β; Vg/dpp; GDF; and Inhibin subgroup patterns. The amino acid sequences defining the finger 1, heel and finger 2 sequence patterns of each subgroup are set forth in Figs. 6A, 6B, and 6C, respectively. In addition, the amino acid sequences defining the entire TGF-β, Vg/dpp, GDF and Inhibin subgroup patterns are set forth in the Sequence Listing as SEQ. ID. Nos. 64, 65, 66, and 67, respectively.

The preferred amino acid sequence patterns for each subgroup, disclosed in Figures 6A, 6B, and 6C, and summarized in Figure 7, enable one skilled in the art to identify alternative amino acids that can be incorporated at specific positions in the finger 1, heel, and finger 2 elements. The amino acids identified in upper case letters in a single letter amino acid code identify conserved amino acids that together are believed to define structural and functional elements of the finger and heel regions. The upper case letter "X" in Figs. 6 and 7 indicates that any naturally occurring amino acid is acceptable at that position. The lower case letter "z" in Figs. 6 and 7 indicates that either a gap or any of the naturally occurring amino acids is acceptable at that position. The lower case letters stand for the amino acids indicated in accordance with the pattern definition table set forth in Figure 8 and identify groups of amino acids which are useful in that location.

In accordance with the amino acid sequence subgroup patterns set forth in Figs. 6-7, it is contemplated, for example, that the skilled artisan can predict that where applicable, one amino acid may be substituted by another without inducing disruptive stereochemical changes within the resulting protein construct. For example, in Fig 6A, in the Vg/dpp subgroup pattern at residue numbers 2 and 3, it is contemplated that either a lysine residue (K) or a arginine residue (R) may be present at this position without affecting the structure of the resulting construct.

Accordingly, the sequence pattern at positions 2 and 3 contain an "n" which, in accordance with Figure 8, defines an amino acid residue selected from the group consisting of lysine or arginine. It is contemplated, therefore, that many synthetic finger 1, finger 2 and heel region amino acid sequences, having 70% homology, preferably 80%, and most preferably at least 90% with the natural regions, can be used to produce conformationally active constructs of the invention.

In accordance with these principles, it is contemplated that one can design a synthetic construct by starting with the amino acid sequence patterns belonging to the TGF-β, Vg/dpp, GDF, or Inhibin subgroup patterns shown in Figs. 6 and 7. Thereafter, by using conventional recombinant or synthetic methodologies a preselected amino acid may be substituted by another as guided by the principles herein and the resulting protein construct tested as provided below.

The TGF-β subgroup pattern, SEQ ID NO: 64, accommodates the homologies shared among members of the TGF-β subgroup identified to date including TGF-β1, TGF-β2, TGF-β3, TGF-β4, and TGF-β5. The generic sequence, shown below, includes both the conserved amino acids (standard three letter code) as well as alternative amino acids (Xaa) present at the variable positions within the sequence and defined by the rules set forth in Fig. 8.

TGF-β Subgroup Pattern

```
Cys Cys Val Arg Pro Leu Tyr Ile Asp Phe Arg Xaa Asp Leu Gly Trp
1               5                   10                  15

Lys Trp Ile His Glu Pro Lys Gly Tyr Xaa Ala Asn Phe Cys Xaa Gly
            20                  25                  30

Xaa Cys Pro Tyr Xaa Trp Ser Xaa Asp Thr Gln Xaa Ser Xaa Val Leu
        35                  40                  45

Xaa Leu Tyr Asn Xaa Xaa Asn Pro Xaa Ala Ser Ala Xaa Pro Cys Cys
    50                  55                  60

Val Pro Gln Xaa Leu Glu Pro Leu Xaa Ile Xaa Tyr Tyr Val Gly Arg
65                  70                  75                  80

Xaa Xaa Lys Val Glu Gln Leu Ser Asn Met Xaa Val Xaa Ser Cys Lys
                85                  90                  95

Cys Ser.
```

Each Xaa can be independently selected from a group of one or more specified amino acids defined as follows, wherein: Xaa12 is Arg or Lys; Xaa26 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa31 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile,Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa33 is Ala, Gly, Pro, Ser, or Thr; Xaa37 is Ile, Leu, Met or Val; Xaa40 isAla, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa44 is His, Phe, Trp or Tyr; Xaa46 is Arg or Lys; Xaa49 is Ala, Gly, Pro, Ser, or Thr; Xaa53 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa54 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa57 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa61 is Ala, Gly, Pro, Ser, or Thr; Xaa68 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa73 is Ala, Gly, Pro, Ser, or Thr; Xaa75 is Ile, Leu, Met or Val; Xaa81 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa82 is Ala, Gly, Pro, Ser, or Thr; Xaa91 is Ile or Val; Xaa93 is Arg or Lys.

The Vg/dpp subgroup pattern, SEQ ID NO: 65, accommodates the homologies shared among members of the Vg/dpp subgroup identified to date including dpp, vg-1, vgr-1, 60A, BMP-2A (BMP-2), Dorsalin, BMP-2B (BMP-4), BMP-3, BMP-5, BMP-6, OP-1 (BMP-7), OP-2 and OP-3. The generic sequence, below, includes both the conserved amino acids (standard three letter code) as well as alternative amino acids (Xaa) present at the variable positions within the sequence and defined by the rules set forth in Fig. 8.

Vg/dpp Subgroup Pattern

```
Cys Xaa Xaa Xaa Xaa Leu Tyr Val Xaa Phe Xaa Asp Xaa Gly Trp Xaa
1               5                   10                  15

Asp Trp Ile Ile Ala Pro Xaa Gly Tyr Xaa Ala Xaa Tyr Cys Xaa Gly
            20                  25                  30

Xaa Cys Xaa Phe Pro Leu Xaa Xaa Xaa Xaa Asn Xaa Thr Asn His Ala
        35                  40                  45

Ile Xaa Gln Thr Leu Val Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro
    50                  55                  60

Lys Xaa Cys Cys Xaa Pro Thr Xaa Leu Xaa Ala Xaa Ser Xaa Leu Tyr
65                  70                  75                  80

Xaa Asp Xaa Xaa Xaa Xaa Xaa Val Xaa Leu Xaa Xaa Tyr Xaa Xaa Met
                85                  90                  95

Xaa Val Xaa Xaa Cys Gly Cys Xaa.
            100
```

Each Xaa can be independently selected from a group of one or more specified amino acids defined as follows, wherein: Xaa2 is Arg or Lys; Xaa3 is Arg or Lys; Xaa4 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa5 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa9 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa11 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa13 is Ile, Leu, Met or Val; Xaa16 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa23 is Arg, Gln, Glu,or Lys; Xaa26 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa28 is Phe, Trp or Tyr; Xaa31 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa33 is Asp or Glu; Xaa35 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa39 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa40 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa41 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa42 is Leu or Met; Xaa44 is Ala, Gly, Pro, Ser, or Thr; Xaa50 is Ile or Val; Xaa55 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa56 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa57 is Ile, Leu, Met or Val; Xaa58 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa59 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond;

Xaa60 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond; Xaa61 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa62 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa63 is Ile or Val; Xaa66 is Ala, Gly, Pro, Ser, or Thr; Xaa69 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa72 is Arg, Gln, Glu,or Lys; Xaa74 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa76 is Ile or Val; Xaa78 is Ile, Leu, Met or Val; Xaa81 is Cys, Ile, Leu, Met, Phe, Trp, Tyr or Val; Xaa83 is Asn, Asp or Glu; Xaa84 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa85 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond; Xaa86 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa87 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa89 is Ile or Val; Xaa91 is Arg or Lys; Xaa92 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa94 is Arg, Gln, Glu,or Lys; Xaa95 is Asn or Asp; Xaa97 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa99 is Arg, Gln, Glu, or Lys; Xaa100 is Ala, Gly, Pro, Ser, or Thr; Xaa104 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr.

The GDF subgroup pattern, SEQ ID NO: 66, accommodates the homologies shared among members of the GDF subgroup identified to date including GDF-1, GDF-3, and GDF-9. The generic sequence, shown below, includes both the conserved amino acids (standard three letter code) as well as alternative amino acids (Xaa) present at the variable positions within the sequence and defined by the rules set forth in Fig. 8.

GDF Subgroup Pattern

```
Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Trp Xaa
1               5                   10                  15

Xaa Trp Xaa Xaa Ala Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Gly
            20                  25                  30

Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
        35                  40                  45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
    50                  55                  60

Pro Xaa Xaa Xaa Xaa Xaa Xaa Cys Val Pro Xaa Xaa Xaa Ser Pro Xaa
65                  70                  75                  80
```

```
Ser Xaa Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Tyr
                85                  90                  95

Glu Asp Met Xaa Xaa Xaa Xaa Cys Xaa Cys Xaa.
            100                 105
```

Each Xaa can be independently selected from a group of one or more specified amino acids defined as follows, wherein: Xaa2 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa3 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa4 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa5 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa6 is Cys, Ile, Leu, Met, Phe, Trp, Tyr or Val; Xaa7 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa8 is Ile, Leu, Met or Val; Xaa9 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa11 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa12 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa13 is Ile, Leu, Met or Val; Xaa14 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa16 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa17 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa19 is Ile or Val; Xaa20 is Ile or Val; Xaa23 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa24 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa25 is Phe, Trp or Tyr; Xaa26 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa27 is Ala, Gly, Pro, Ser, or Thr; Xaa28 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa29 is Phe, Trp or Tyr; Xaa31 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa33 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa35 is Ala, Gly, Pro, Ser, or Thr; Xaa36 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa37 is Ala, Gly, Pro, Ser, or Thr; Xaa38 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa39 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa40 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa41 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa42 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa43 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond; Xaa44 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr,

Val or a peptide bond; Xaa45 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond; Xaa46 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond; Xaa47 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa48 is Ala, Gly, Pro, Ser, or Thr; Xaa49 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa50 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa51 is His, Phe, Trp or Tyr; Xaa52 is Ala, Gly, Pro, Ser, or Thr; Xaa53 is Cys, Ile, Leu, Met, Phe, Trp, Tyr or Val; Xaa54 is Ile, Leu, Met or Val; Xaa55 is Arg, Gln, Glu,or Lys; Xaa56 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa57 is Ile, Leu, Met or Val; Xaa58 is Ile, Leu, Met or Val; Xaa59 is His, Phe, Trp or Tyr; Xaa60 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa61 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa62 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa63 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond; Xaa64 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond; Xaa66 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa67 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa68 is Ala, Gly, Pro, Ser, or Thr; Xaa69 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa70 is Ala, Gly, Pro, Ser, or Thr; Xaa71 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa75 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa76 is Arg or Lys; Xaa77 is Cys, Ile, Leu, Met, Phe, Trp, Tyr or Val; Xaa80 is Ile, Leu, Met or Val; Xaa82 is Ile, Leu, Met or Val; Xaa84 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa85 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa86 is Asp or Glu; Xaa87 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa88 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa89 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa90 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or

Thr; Xaa91 is Ile or Val; Xaa92 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa93 is Cys, Ile, Leu, Met, Phe, Trp, Tyr or Val; Xaa94 is Arg or Lys; Xaa95 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa100 is Ile or Val; Xaa101 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa102 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa103 is Arg, Gln, Glu,or Lys; Xaa105 is Ala, Gly, Pro, Ser, or Thr; Xaa107 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val.

The Inhibin subgroup pattern, SEQ ID NO: 67, accommodates the homologies shared among members of the Inhibin subgroup identified to date including Inhibin α, Inhibin βA and Inhibin βB. The generic sequence, shown below, includes both the conserved amino acids (standard three letter code) as well as alternative amino acids (Xaa) present at the variable positions within the sequence and defined by the rules set forth in Fig. 8.

Inhibin Subgroup pattern

```
Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Gly Trp Xaa
1               5                   10                  15

Xaa Trp Ile Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Tyr Cys Xaa Gly
            20                  25                  30

Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
        35                  40                  45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
    50                  55                  60

Xaa Xaa Xaa Xaa Xaa Cys Cys Xaa Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa
65                  70                  75                  80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asp Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                85                  90                  95

Xaa Xaa Xaa Asn Xaa Xaa Xaa Xaa Xaa Cys Xaa Cys Xaa.
            100                 105
```

Each Xaa can be independently selected from a group of one or more specified amino acids defined as follows, wherein: Xaa2 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa3 is Arg or Lys; Xaa4 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa5 is Ala,

Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa6 is Cys, Ile, Leu, Met, Phe, Trp, Tyr or Val; Xaa7 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa8 is Ile or Val; Xaa9 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa11 is Arg, Gln, Glu,or Lys; Xaa12 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa13 is Ile, Leu, Met or Val; Xaa16 is Asn, Asp or Glu; Xaa17 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa20 is Ile or Val; Xaa21 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa23 is Ala, Gly, Pro, Ser, or Thr; Xaa24 is Ala, Gly, Pro, Ser, or Thr; Xaa25 is Phe, Trp or Tyr; Xaa26 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa27 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa28 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa31 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa33 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa35 is Ala, Gly, Pro, Ser, or Thr; Xaa36 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa37 is His, Phe, Trp or Tyr; Xaa38 is Ile, Leu, Met or Val; Xaa39 is Ala, Gly, Pro, Ser, or Thr; Xaa40 is Ala, Gly, Pro, Ser, or Thr; Xaa41 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa42 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa43 is Ala, Gly, Pro, Ser, or Thr; Xaa44 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa45 is Ala, Gly, Pro, Ser, or Thr; Xaa46 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa47 is Ala, Gly, Pro, Ser, or Thr; Xaa48 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa49 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa50 is Ala, Gly, Pro, Ser, or Thr; Xaa51 is Ala, Gly, Pro, Ser, or Thr; Xaa52 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa53 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa54 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa55 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa56 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa57

is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa58 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa59 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa60 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond; Xaa61 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond; Xaa62 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond; Xaa63 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond; Xaa64 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa65 is Ala, Gly, Pro, Ser, or Thr; Xaa66 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa67 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa68 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa69 is Ala, Gly, Pro, Ser, or Thr; Xaa72 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa73 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond; Xaa74 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond; Xaa76 is Ala, Gly, Pro, Ser, or Thr; Xaa77 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa78 is Leu or Met; Xaa79 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa80 is Ala, Gly, Pro, Ser, or Thr; Xaa81 is Leu or Met; Xaa82 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa83 is Ile, Leu, Met or Val; Xaa84 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa85 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa86 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa87 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa89 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa90 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr, Val or a peptide bond; Xaa91 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa92 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa93 is Cys, Ile, Leu, Met, Phe, Trp, Tyr or Val; Xaa94 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe,

Pro, Ser, Thr, Trp, Tyr or Val; Xaa95 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa96 is Arg, Gln, Glu,or Lys; Xaa97 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa98 is Ile or Val; Xaa99 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa101 is Leu or Met; Xaa102 is Ile, Leu, Met or Val; Xaa103 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val; Xaa104 is Gln or Glu; Xaa105 is Arg, Asn, Asp, Gln, Glu, His, Lys, Ser or Thr; Xaa107 is Ala or Gly; Xaa109 is Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr or Val.

B. Recombinant Production Considerations

1. Design and Production of the Protein and DNA Constructs of the Present Invention

As mentioned above, the constructs of the invention can be manufactured by using conventional recombinant DNA methodologies well known and thoroughly documented in the art, as well as by using well-known biosynthetic and chemosynthetic methodologies using routine peptide or nucleotide chemistries and automated peptide or nucleotide synthesizers. Such routine methodologies are described for example in the following publications, the teachings of which are incorporated by reference herein: Hilvert, 1 Chem. Biol. 201-3 (1994); Muir et al., 95 Proc. Natl. Acad. Sci. USA 6705-10 (1998); Wallace, 6 Curr. Opin. Biotechnol. 403-10 (1995); Miranda et al., 96 Proc. Natl. Acad. Sci. USA 1181-86 (1999); Liu et al., 91 Proc. Natl. Acad. Sci. USA 6584-88 (1994). Suitable for use in the present invention are naturally-occurring amino acids and nucleotides; non-naturally occurring amino acids and nucleotides; modified or unusual amino acids; modified bases; amino acid sequences that contain post-translationally modified amino acids and/or modified linkages, cross-links and end caps, non-peptidyl bonds, etc.; and, further including without limitation, those moieties disclosed in the World Intellectual Property Organization (WIPO) Handbook on Industrial Property Information and Documentation, Standard St. 25 (1998) including Tables 1 through 6 in Appendix 2, herein incorporated by reference. Equivalents of the foregoing will be appreciated by the skilled artisan relying only on routine experimentation together with the knowledge of the art.

For example, the contemplated DNA constructs may be manufactured by the assembly of synthetic nucleotide sequences and/or joining DNA restriction fragments to produce a synthetic DNA molecule. The DNA molecules then are ligated into an expression vehicle, for example an expression plasmid, and transfected into an appropriate host cell, for example *E. coli*. The contemplated protein construct encoded by the DNA molecule then is expressed, purified, refolded, tested *in vitro* for certain attributes, e.g., binding activity with a receptor having binding affinity for the template TGF-β superfamily member, and subsequently tested to assess whether the biosynthetic construct mimics other preferred attributes of the template superfamily member.

Alternatively, a library of synthetic DNA constructs can be prepared simultaneously for example, by the assembly of synthetic nucleotide sequences that differ in nucleotide composition in a preselected region. For example, it is contemplated that during production of a construct based upon a specific TGF-β superfamily member, the artisan can choose appropriate finger and heel regions for such a superfamily member (for example from Figs. 6-7). Once the appropriate finger and heel regions have been selected, the artisan then can produce synthetic DNA encoding these regions. For example, if a plurality of DNA molecules encoding different linker sequences are included into a ligation reaction containing DNA molecules encoding finger and heel sequences, by judicious choice of appropriate restriction sites and reaction conditions, the artisan may produce a library of DNA constructs wherein each of the DNA constructs encode finger and heel regions but connected by different linker sequences. The resulting DNAs then are ligated into a suitable expression vehicle, i.e., a plasmid useful in the preparation of a phage display library, transfected into a host cell, and the polypeptides encoded by the synthetic DNAs expressed to generate a pool of candidate proteins. The pool of candidate proteins subsequently can be screened to identify specific proteins having the desired binding affinity and/or selectivity for a pre-selected receptor.

Screening can be performed by passing a solution comprising the candidate proteins through a chromatographic column containing surface immobilized receptor. Then lead proteins with the desired binding specificity are eluted, for example by means of a salt gradient and/or a concentration gradient of the template TGF-β superfamily member. Nucleotide sequences encoding such proteins subsequently can be isolated and characterized. Once the appropriate nucleotide sequences have been identified, the proteins subsequently can be produced, either by

conventional recombinant DNA or peptide synthesis methodologies, in quantities sufficient to test whether the particular construct mimics the activity of the template TGF-β superfamily member.

It is contemplated that, which ever approach is adopted to produce DNA molecules encoding constructs of the invention, the tertiary structure of the preferred proteins can subsequently be modulated in order to optimize binding and/or biological activity by, for example, by a combination of nucleotide mutagenesis methodologies aided by the principles described herein and phage display methodologies. Accordingly, an artisan can produce and test simultaneously large numbers of such proteins.

(a) Gene and Protein Synthesis

The processes for manipulating, amplifying, and recombining DNA which encode amino acid sequences of interest generally are well known in the art, and therefore, are not described in detail herein. Methods of identifying and isolating genes encoding members of the TGF-β superfamily also are well understood, and are described in the patent and other literature.

Briefly, the construction of DNAs encoding the biosynthetic constructs disclosed herein is performed using known techniques involving the use of various restriction enzymes which make sequence specific cuts in DNA to produce blunt ends or cohesive ends, DNA ligases, techniques enabling enzymatic addition of sticky ends to blunt-ended DNA, construction of synthetic DNAs by assembly of short or medium length oligonucleotides, cDNA synthesis techniques, polymerase chain reaction (PCR) techniques for amplifying appropriate nucleic acid sequences from libraries, and synthetic probes for isolating genes of members of the TGF-β superfamily and their cognate receptors. Various promoter sequences from bacteria, mammals, or insects to name a few, and other regulatory DNA sequences used in achieving expression, and various types of host cells are also known and available. Conventional transfection techniques, and equally conventional techniques for cloning and subcloning DNA are useful in the practice of this invention and known to those skilled in the art. Various types of vectors may be used such as plasmids and viruses including animal viruses and bacteriophages. The vectors may exploit various marker genes which impart to a successfully transfected cell a detectable phenotypic property that can be used to identify which of a family of clones has successfully incorporated the recombinant DNA of the vector.

One method for obtaining DNA encoding the biosynthetic constructs disclosed herein is by assembly of synthetic oligonucleotides produced in a conventional, automated, oligonucleotide synthesizer followed by ligation with appropriate ligases. For example, overlapping, complementary DNA fragments may be synthesized using phosphoramidite chemistry, with end segments left unphosphorylated to prevent polymerization during ligation. One end of the synthetic DNA is left with a "sticky end" corresponding to the site of action of a particular restriction endonuclease, and the other end is left with an end corresponding to the site of action of another restriction endonuclease. The complementary DNA fragments are ligated together to produce a synthetic DNA construct.

Alternatively nucleic acid strands encoding finger 1, finger 2 and heel regions can be isolated from libraries of nucleic acids, for example, by colony hybridization procedures such as those described in Sambrook et al. eds. (1989) "Molecular Cloning", Coldspring Harbor Laboratories Press, NY, and/or by PCR amplification methodologies, such as those disclosed in Innis et al. (1990) "PCR Protocols, A guide to methods and applications", Academic Press. The nucleic acids encoding the finger and heel regions then are joined together to produce a synthetic DNA encoding the biosynthetic single-chain morphon construct of interest.

It is appreciated, however, that a library of DNA constructs encoding a plurality thereof can be produced simultaneously by standard recombinant DNA methodologies, such as the ones, described above, For example, the skilled artisan by the use of cassette mutagenesis or oligonucleotide directed mutagenesis can produce, for example, a series of DNA constructs each of which contain different DNA sequences within a predefined location, e.g., within a DNA cassette encoding a linker sequence. The resulting library of DNA constructs subsequently can be expressed, for example, in a phage or viral display library or a eukaryotic cell line (e.g., CHO cell line); and any protein constructs that binds to a specific receptor may be isolated by affinity purification, e.g., using a chromatographic column comprising surface immobilized receptor. Once molecules that bind the preselected receptor have been isolated, their binding and agonist properties can be modulated using empirical refinement techniques.

Methods of mutagenesis of proteins and nucleic acids are well known and well described in the art. See, e.g., Sambrook et al., (1990) Molecular Cloning: A Laboratory Manual., 2d ed. (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press). Useful methods include PCR

(overlap extension, see, e.g., PCR Primer (Dieffenbach and Dveksler, eds., Cold Spring Harbor Press, Cold Spring Harbor, NY, 1995, pp. 603-611)); cassette mutagenesis and single-stranded mutagenesis following the method of Kunkel. It will be appreciated by the artisan that any suitable method of mutagenesis can be utilized and the mutagenesis method is not considered a material aspect of the invention. The nucleotide codons competent to encode amino acids, including arginine (Arg), glutamic acid (Glu)and aspartic acid (Asp) also are well known and described in the art. See, for example, Lehninger, Biochemistry (Worth Publishers, N.Y., N.Y.). Standard codons encoding arginine, glutamic acid and aspartic acid are: Arg: CGU, CGC, CGA, CGG, AGA, AGG; Glu: GAA, GAG; and Asp: GAU, GAC. Chimeric constructs of the invention can readily be constructed by aligning the nucleic acid sequences or domains to be switched, and identifying compatible splice sites and/or constructing suitable crossover sequences using PCR overlap extension.

(b) Protein Expression

One of the advantages of the present invention relates to protein expression and production, especially commercially significant quantities of protein. As disclosed herein, chimeric proteins can be endowed with certain preferred attributes such as enhanced/improved refolding attributes. Moreover, chimeric proteins with enhanced/improved refolding attributes can be designed and expressed in bacterial cells that could not otherwise be properly expressed in mammalian cells. It is contemplated that such enhanced refolding will improve recovery and commercial production of active TGF-β superfamily proteins. It is also now appreciated that mammalian cells can identify misfolded proteins and have cellular processes which permit such proteins to be recycled and then refolded properly. Such proteins will be discarded, however, if they fail to refold properly ultimately. Thus the less time a cell spends re-cycling poor refolder proteins the more time available to produce usable refolded proteins. See, for example, Trombetta et al., 8 Curr. Opin. Struct. Biol. 587-92 (1998).

If a single DNA construct of interest has been synthesized, it can be integrated into an expression vector and transfected into an appropriate host cell for protein expression. Useful prokaryotic host cells include, but are not limited to, *E. coli*, and *B. Subtilis*. Useful eukaryotic host cells include, but are not limited to, yeast cells, insect cells, myeloma cells, fibroblast 3T3

cells, monkey kidney or COS cells, chinese hamster ovary (CHO) cells, mink-lung epithelial cells, human foreskin fibroblast cells, human glioblastoma cells, and teratocarcinoma cells. Alternatively, synthetic genes may be expressed in a cell-free system such as the rabbit reticulocyte lysate system.

The vector additionally can include various sequences to promote correct expression of the recombinant protein, including transcriptional promoter and termination sequences, enhancer sequences, preferred ribosome binding site sequences, preferred mRNA leader sequences, preferred protein processing sequences, preferred signal sequences for protein secretion, and the like. The DNA sequence encoding the gene of interest also can be manipulated to remove potentially inhibiting sequences or to minimize unwanted secondary structure formation. The modified proteins of the present invention also may be expressed as fusion proteins. After being translated, the protein can be purified from the cells themselves or recovered from the culture medium and then cleaved at a specific protease site if so desired.

For example, if the gene is to be expressed in *E. coli*, it is cloned into an appropriate expression vector. This can be accomplished by positioning the engineered gene downstream of a promoter sequence such as Trp or Tac, and/or a gene coding for a leader peptide such as fragment B of protein A (FB). During expression, the resulting fusion proteins accumulate in refractile bodies in the cytoplasm of the cells, and may be harvested after disruption of the cells by French press or sonication. The isolated refractile bodies then are solubilized, and the expressed proteins refolded as taught herein.

If desired, expression of the engineered genes in eukaryotic cells can be achieved. This requires cells and cell lines that are easy to transfect, are capable of stably maintaining foreign DNA with an unrearranged sequence, and which have the necessary cellular components for efficient transcription, translation, post-translation modification, and secretion of the protein. In addition, a suitable vector carrying the gene of interest also is necessary. DNA vector design for transfection into mammalian cells should include appropriate sequences to promote expression of the gene of interest as described herein, including appropriate transcription initiation, termination, and enhancer sequences, as well as sequences that enhance translation efficiency, such as the Kozak consensus sequence. Preferred DNA vectors also include a marker gene and means for amplifying the copy number of the gene of interest. A detailed review of the state of

the art of the production of foreign proteins in mammalian cells, including useful cells, protein expression-promoting sequences, marker genes, and gene amplification methods, is disclosed in Bendig (1988) Genetic Engineering 7:91-127.

The best characterized transcription promoters useful for expressing a foreign gene in a particular mammalian cell are the SV40 early promoter, the adenovirus promoter (AdMLP), the mouse metallothionein-I promoter (mMT-I), the Rous sarcoma virus (RSV) long terminal repeat (LTR), the mouse mammary tumor virus long terminal repeat (MMTV-LTR), and the human cytomegalovirus major intermediate-early promoter (hCMV). The DNA sequences for all of these promoters are known in the art and are available commercially.

The use of a selectable DHFR gene in a dhfr$^-$ cell line is a well characterized method useful in the amplification of genes in mammalian cell systems. Briefly, the DHFR gene is provided on the vector carrying the gene of interest, and addition of increasing concentrations of the cytotoxic drug methotrexate, which is metabolized by DHFR, leads to amplification of the DHFR gene copy number, as well as that of the associated gene of interest. DHFR as a selectable, amplifiable marker gene in transfected chinese hamster ovary cell lines (CHO cells) is particularly well characterized in the art. Other useful amplifiable marker genes include the adenosine deaminase (ADA) and glutamine synthetase (GS) genes.

The choice of cells/cell lines is also important and depends on the needs of the experimenter. COS cells provide high levels of transient gene expression, providing a useful means for rapidly screening the biosynthetic constructs of the invention. COS cells typically are transfected with a simian virus 40 (SV40) vector carrying the gene of interest. The transfected COS cells eventually die, thus preventing the long term production of the desired protein product. However, transient expression does not require the time consuming process required for the development of a stable cell line, and thus provides a useful technique for testing preliminary constructs for binding activity.

A particular advantage of the methods and biosynthetic recombinant or chemosynthetic proteins of the instant invention is that they are well suited for use in yeast and bacterial cell systems such as *E. coli* and other systems where overexpression results in formation of inclusion bodies from which the proteins must be resolubilized and refolded *in vitro*. Even though TGF-β superfamily member protein mutants can be made in mammalian cells, some designed protein

constructs and mutants appear to be unstable in mammalian cells and can only be made, at the current state of the art, by using bacterial cells and refolding the mutants *in vitro*. Detailed descriptions of the proteins useful in the practice of this invention, including how to make, use and test them for osteogenic activity, are disclosed in numerous publications, including US Patent Nos. 5,266,683 and 5,011,691, the disclosures of which are incorporated by reference herein, as well as in any of the publications recited herein, the disclosures of which are incorporated herein by reference.

Briefly, the chimeric forms of the present invention can be expressed and produced in bacteria and yeast using standard, well-known methods. Full-length mature forms or shorter sequences defining only the C-terminal seven cysteine domain can be provided to the host cell. It may be preferred to modify the N-terminal sequences to optimize bacterial expression. For example, the preferred form of native OP-1 for bacterial expression is the sequence encoding the mature, active sequence (residues 293-431 of SEQ No. 39) or a fragment thereof encoding the C-terminal seven cysteine domain (e.g., residues 330-431 of SEQ ID NO: 39). A methionine can be introduced at position 293, replacing the native serine residue, or it can precede this serine residue. Alternatively, a methionine can be introduced anywhere within the first thirty-six residues of the natural sequence (residues 293-329). The DNA sequence further can be modified at its N-terminus to improve purification, for example, by adding a "hexa-His" tail to assist purification on an IMAC column; or by using a FB leader sequence, which facilitates purification on an IgG/column. These and other methods are well described and well known in the art. Other bacterial species and/or proteins may require or benefit from analogous modifications to optimize the yield of the mutant BMP obtained therefrom. Such modifications are well within the level of ordinary skill in the art and are not considered material aspects of the invention.

The synthetic nucleic acids preferably are inserted into a vector suitable for overexpression in the host cell of choice. Any expression vector can be used, so long as it is capable of directing the expression of a heterologous protein such as a BMP in the host cell of choice. Useful vectors include plasmids, phagemids, mini chromosomes and YACs, to name a few. Other vector systems are well known and characterized in the art. The vector typically includes a replicon, one or more selectable marker gene sequences, and means for maintaining a high copy number of the vector in the host cell. Well known selectable marker genes include

antibiotics like ampicillin, tetracycline and the like, as well as resistance to heavy metals. Useful selectable marker genes for use in yeast cells include the URA3, LEU2, HIS3 or TRP1 gene for use with an auxotrophic yeast mutant host. In addition, the vector also includes a suitable promoter sequence for expressing the gene of interest and which may or may not be inducible, as desired, as well as useful transcription and translation initiation sites, terminators, and other sequences that can maximize transcription and translation of the gene of interest. Well characterized promotors particularly useful in bacterial cells include the lac, tac, trp, and tpp promoters, to name a few. Promoters useful in yeast include ADHI, ADHII, or PHO5 promoter, for example.

(c) Refolding Considerations

Proteins produced by *E-coli* for example, are first isolated from inclusion bodies, and then solubilized using a denaturant or chaotropic agent such as guanidine HCl or urea, preferably in the range of about 4-9 M and at an elevated temperature (e.g., 25-37° C) and/or basic pH (8-10). Alternatively, the proteins can be solubilized by acidification, e.g., with acetic acid or trifluoroacetic acid, generally at a pH in the range of 1-4. Preferably, a reducing agent such as β-mercaptoethanol or dithiothreitol (DTT) is used in conjunction with the solubilizing agent. The solubilized heterologous protein can be purified further from solubilizing chaotropes by dialysis and/or by known chromatographic methods such as size exclusion chromatography, ion exchange chromatography, or reverse phase high performance liquid chromatography (RP-HPLC), for example.

In accordance with the present invention, a solubilized chimera can be refolded as follows. The dissolved protein is diluted in a refolding medium, typically a Tris-buffered medium having a pH in the range of about pH 5.0-10.0, preferably in the range of about pH 6-9 and one which includes a detergent and/or chaotropic agent. Useful commercially available detergents can be ionic, nonionic, or zwitterionic, such as NP40 (Nonidet 40), CHAPS (such as 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfate, digitonin, deoxycholate, or N-octyl glucoside. Useful chaotropic agents include guanidine, urea, or arginine. Preferably the chaotropic agent is present at a concentration in the range of about 0.1-10M, preferably in the range of about 0.5-4M. When CHAPS is the detergent, it preferably comprises about 0.5-5% of

the solution, more preferably about 1-3% of the solution. Preferably the solution also includes a suitable redox system such as the oxidized and reduced forms of glutathione, DTT, β-mercaptoethanol, β-mercaptomethanol, cysteine or cystamine, to name a few. Preferably, the redox systems are present at ratios of reductant to oxidant in the range of about 1:1 to about 5:1. When the glutathione redox system is used, the ratio of reduced glutathione to oxidized glutathione is preferably is in the range of about 0.5 to 5; more preferably 1 to 1; and most preferably 2 to 1 of reduced form to oxidized form. Preferably the buffer also contains a salt, typically NaCl, present in the range of about 0.25M –2.5 M, preferably in the range of about 0.5-1.5M, most preferably in the range of about 1M. One skilled in the art will recognize that the above conditions and media may be optimized using no more than routine experimentation. Such variations and modifications are within the scope of the present invention.

Preferably the protein concentration for a given refolding reaction is in the range of about 0.001-1.0 mg/ml, more preferably it is in the range of about 0.05-0.25 mg/ml, most preferably in the range of about 0.075-0.125 mg/ml. As will be appreciated by the skilled artisan, higher concentrations tend to produce more aggregates. Where heterodimers are to be produced (for example an OP-1/BMP-2 or BMP-2/BMP-6 heterodimer) preferably the individual proteins are provided to the refolding buffer in equal amounts.

Typically, the refolding reaction takes place at a temperature range from about 4°C to about 25°C. More preferably, the refolding reaction is performed at 4 °C, and allowed to go to completion. Refolding typically is complete in about one to seven days, generally within 16-72 hours or 24-48 hours, depending on the protein. As will be appreciated by the skilled artisan, rates of refolding can vary by protein, and longer and shorter refolding times are contemplated and within the scope of the present invention. As used herein, a "good refolder" protein is one where at least approximately 10% of the protein, more preferably at least approximately 20% and even more preferably at least approximately 25%, and most preferably greater than 25%, is refolded following a folding reaction when compared to the total protein in the refolding reaction, as measured by any of the refolding assays described herein and without requiring further purification. For example, native BMPs that are considered in the art to be "good refolder" proteins include BMP-2, CDMP-1, CDMP-2 and CDMP-3. BMP-3 also refolds reasonably well. In contrast, a "poor refolder" protein yields less than 1% of optimally-folded

protein. When attributes are optimally mixed-and-matched as disclosed herein, certain embodiments can yield more than approximately 25% protein in an optimal re-folded state.

2. Summary of Assays for Screening, Binding and Testing of Biological Activity

Irrespective of which protein expression, harvesting, and, folding methodologies are used, certain of the chimeric proteins can bind preferentially to a preselected receptor and can now be identified using standard methodologies, i.e., ligand/receptor binding assays, well known, and thoroughly documented in the art. See for example: Legerski et al. (1992) Biochem. Biophys. Res. Comm. 183: 672-679; Frakar et al. (1978) Biochem. Biophys. Res. Comm 80:849-857; Chio et al. (1990) Nature 343: 266-269; Dahlman et al. (1988) Biochem 27: 1813-1817; Strader et al. (1989) J. Biol. Chem. 264: 13572-13578; and D'Dowd et al. (1988) J. Biol. Chem. 263: 15985-15992.

Typically, in a ligand/receptor binding assay, the native or parent TGF-β superfamily member of interest having a known, quantifiable affinity for a preselected receptor is labeled with a detectable moiety, for example, a radiolabel, a chromogenic label, or a fluorogenic label. Aliquots of purified receptor, receptor binding domain fragments, or cells expressing the receptor of interest on their surface are incubated with the labeled TGF-β superfamily member in the presence of various concentrations of the unlabeled chimeric protein. The relative binding affinity of a candidate chimera may be measured by quantitating the ability of the chimera to inhibit the binding of the labeled TGF-β superfamily member with the receptor. In performing the assay, fixed concentrations of the receptor and the TGF-β superfamily member are incubated in the presence and absence of unlabeled chimera. Sensitivity may be increased by pre-incubating the receptor with the chimera before adding the labeled template TGF-β superfamily member. After the labeled competitor has been added, sufficient time is allowed for adequate competitor binding, and then free and bound labeled superfamily members are separated from one another, and one or the other measured.

Labels useful in the practice of the screening procedures include radioactive labels, i.e., ^{125}I, ^{131}I, ^{111}In or ^{77}Br, chromogenic labels, spectroscopic labels such as those disclosed in Haughland (1994) "Handbook of Fluorescent and Research Chemicals 5 ed." by Molecular

Probes, Inc., Eugene, OR, or conjugated enzymes having high turnover rates, i.e., horseradish peroxidase, alkaline phosphatase, or β-galactosidase, used in combination with chemiluminescent or fluorogenic substrates.

The biological activity, namely the agonist or antagonist properties of the resulting chimeric constructs can subsequently be characterized using conventional *in vivo* and *in vitro* assays that have been developed to measure the biological activity of any TGF-β superfamily member. It is appreciated, however, the type of assay used will depend on the TGF-β superfamily member upon which the chimera is based. For example, chimeric constructs based upon natural OP-1 protein may be assayed using any of the biological assays that have been developed to date for measuring OP-1 activity, see, for example, the exemplification's set forth below.

Optimally refolded dimeric proteins readily can be assessed using any of a number of well known and well characterized assays. In particular, any one or more of three assays, all well known and well described in the art, and further described below can be used to advantage.

Useful refolding assays include one or more of the following. First, the presence of dimers can be detected visually either by standard SDS-PAGE in the absence of a reducing agent such as DTT or by HPLC (e.g., C18 reverse phase HPLC). Dimeric proteins of the present invention have an apparent molecular weight in the range about 28-36 kDa, as compared to monomeric subunits, which have an apparent molecular weight of about 14-18 kDa. The dimeric protein can readily be visualized on an electrophoresis gel by comparison to commercially available molecular weight standards. The dimeric protein also elutes from a C18 RP HPLC (45-50% acetonitrile: 0.1%TFA) at about 19 minutes (mammalian produced hOP-1 elutes at 18.95 minutes). A second assay evaluates the presence of dimer by its ability to bind to hydroxyapatite. Optimally-folded dimer binds a hydroxyapatite column well in pH 7, 10mM phosphate, 6M urea and 0.1-0.2M NaCl (dimer elutes at 0.25 M NaCl) as compared to monomer, which does not bind substantially at those concentrations (monomer elutes at 0.1M NaCl). A third assay evaluates the presence of dimer by the protein's resistant to trypsin or pepsin digestion. The folded dimeric species is substantially resistant to both enzymes, particularly trypsin, which cleaves only a small portion of the N-terminus of the mature protein, leaving a biologically active dimeric species only slightly smaller in size than the untreated dimer (each

monomer is 22 amino acids smaller after trypsin cleavage). By contrast, the monomers and mis-folded dimers are substantially degraded. In the assay, the protein is subjected to an enzyme digest using standard conditions, e.g., digestion in a standard buffer such as 50mM Tris buffer, pH 8, containing 4 M urea, 100 mM NaCl, 0.3% Tween-80 and 20 mM methylamine. Digestion is allowed to occur at 37°C for on the order of 16 hours, and the product visualized by any suitable means, preferably SDS-PAGE.

The biological activity of the refolded chimeric proteins, for example BMPs, readily can be assessed by any of a number of means as described below. For example, the protein's ability to induce endochondral bone formation can be evaluated using the well characterized rat subcutaneous bone assay, described in detail below. In the assay bone formation is measured by histology, as well as by alkaline phosphatase and/or osteoclacin production. In addition, osteogenic proteins having high specific bone forming activity, such as OP-1, BMP-2, BMP-4, BMP-5 and BMP-6, also induce alkaline phosphatase activity in an in vitro rat osteoblast or osteosarcoma cell-based assay. Such assays are well described in the art and are detailed herein below. See, for example, Sabokdar et al. (1994) Bone and Mineral 27:57-67.; Knutsen et al. (1993) Biochem. Biophys. Res. Commun. 194:1352-1358; and Maliakal et al. (1994) Growth Factors 1:227-234). By contrast, osteogenic proteins having low specific bone forming activity, such as CDMP-1 and CDMP-2, for example, do not induce alkaline phosphatase activity in the cell based osteoblast assay. The assay thus provides a ready method for evaluating biological activity of BMP mutants. For example, CDMP-1, CDMP-2 and CMDP-3 all are competent to induce bone formation, although with a lower specific activity than BMP-2, BMP-4, BMP-5, BMP-6 or OP-1. Conversely, BMP-2, BMP-4, BMP-5, BMP-6 and OP-1 all can induce articular cartilage formation, albeit with a lower specific activity than CDMP-1, CDMP-2 or CDMP-3. Accordingly, a CDMP mutant, designed and described herein to be a mutant competent to induce alkaline phosphatase activity in the cell-based assay, is expected to demonstrate a higher specific bone forming activity in the rat animal bioassay. Similarly, an OP-1 mutant designed and described herein to be containing a substitution present in a corresponding position of a CDMP-1, CDMP-2 or CDMP-3 protein, is competent to induce bone in the rat assay but not to induce alkaline phosphatase activity in the cell based assay, and thus is expected to have a higher specific articular cartilage inducing activity in an *in vivo* articular cartilage assay. As described

herein below, a suitable *in vitro* assay for CDMP activity utilizes mouse embryonic osteoprogenitor or carcinoma cells, such as ATDC5 cells. See Example 6, below.

TGF-ß activity can be readily evaluated by the protein's ability to inhibit epithelial cell growth. A useful, well characterized *in vitro* assay utilizes mink lung cells or melanoma cells. See Example 7. Other assays for other members of the TGF-ß superfamily are well described in the literature and can be performed without undue experimentation.

3. Formulation and Bioactivity

The resulting chimeric proteins can be provided to an individual as part of a therapy to enhance, inhibit, or otherwise modulate *in vivo* events, such as but not limited to, the binding interaction between a TGF-β superfamily member and one or more of its cognate receptors. The constructs may be formulated in a pharmaceutical composition, as described below, and may be administered in morphogenic effective amounts by any suitable means, preferably directly or systematically, e.g., parentally or orally. Resulting DNA constructs encoding preferred chimeric proteins can also be administered directly to a recipient for gene therapeutic purposes; such DNAs can be administered with or without carrier components, or with or without matrix components. Alternatively, cells transferred with such DNA constructs can be implanted in a recipient. Such materials and methods are well-known in the art.

Where any of the constructs disclosed here are to be provided directly (e.g., locally, as by injection, to a desired tissue site), or parentally, such as by intravenous, subcutaneous, intramuscular, intraorbital, ophthalmic, intraventricular, intracranial, intracapsular, intraspinal, intracisternal, intraperitoneal, buccal, rectal, vaginal, intranasal or by aerosol administration, the therapeutic composition preferably comprises part of an aqueous solution. The solution preferably is physiologically acceptable so that in addition to delivery of the desired construct to the patient, the solution does not otherwise adversely affect the patient's electrolyte and volume balance. The aqueous medium for the therapeutic molecule thus may comprise, for example, normal physiological saline (0.9% NaCl, 0.15M), pH 7-7.4 or other pharmaceutically acceptable salts thereof.

Useful solutions for oral or parenteral administration may be prepared by any of the methods well known in the pharmaceutical art, described, for example, in Remington's

Pharmaceutical Sciences, (Gennaro, A., ed.), Mack Pub., 1990. Formulations may include, for example, polyalkylene glycols such as polyethylene glycol, oils of vegetable origin, hydrogenated naphthalenes, and the like. Formulations for direct administration, in particular, may include glycerol and other compositions of high viscosity. Biocompatible, preferably bioresorbable polymers, including, for example, hyaluronic acid, collagen, tricalcium phosphate, polybutyrate, polylactide, polyglycolide and lactide/glycolide copolymers, may be useful excipients to control the release of the morphogen *in vivo*.

Other potentially useful parenteral delivery systems for these therapeutic molecules include ethylene-vinyl acetate copolymer particles, osmotic pumps, implantable infusion systems, and liposomes. Formulations for inhalation administration may contain as excipients, for example, lactose, or may be aqueous solutions containing, for example, polyoxyethylene-9-lauryl ether, glycocholate and deoxycholate, or oily solutions for administration in the form of nasal drops, or as a gel to be applied intranasally.

Finally, therapeutic molecules may be administered alone or in combination with other molecules known to effect tissue morphogenesis, i.e., molecules capable of tissue repair and regeneration and/or inhibiting inflammation. Examples of useful cofactors for stimulating bone tissue growth in osteoporotic individuals, for example, include but are not limited to, vitamin D_3, calcitonin, prostaglandins, parathyroid hormone, dexamethasone, estrogen and IGF-I or IGF-II. Useful cofactors for nerve tissue repair and regeneration may include nerve growth factors. Other useful cofactors include symptom-alleviating cofactors, including antiseptics, antibiotics, antiviral and antifungal agents and analgesics and anesthetics.

Therapeutic molecules further can be formulated into pharmaceutical compositions by admixture with pharmaceutically acceptable nontoxic excipients and carriers. As noted above, such compositions may be prepared for parenteral administration, particularly in the form of liquid solutions or suspensions; for oral administration, particularly in the form of tablets or capsules; or intranasally, particularly in the form of powders, nasal drops or aerosols. Where adhesion to a tissue surface is desired the composition may include the biosynthetic construct dispersed in a fibrinogen-thrombin composition or other bioadhesive such as is disclosed, for example in PCT US91/09275, the disclosure of which is incorporated herein by reference. The composition then may be painted, sprayed or otherwise applied to the desired tissue surface.

The compositions can be formulated for parenteral or oral administration to humans or other mammals in therapeutically effective amounts, e.g., amounts which provide appropriate concentrations of the morphon to target tissue for a time sufficient to induce the desired effect.

Where the therapeutic molecule comprises part of a tissue or organ preservation solution, any commercially available preservation solution may be used to advantage. For example, useful solutions known in the art include Collins solution, Wisconsin solution, Belzer solution, Eurocollins solution and lactated Ringer's solution. A detailed description of preservation solutions and useful components may be found, for example, in U.S. Patent No. 5,002,965, the disclosure of which is incorporated herein by reference.

It is contemplated that some of the protein constructs, for example those based upon members of the Vg/dpp subgroup will also exhibit high levels of activity *in vivo* when combined with a matrix. See for example, U.S. Patent No. 5,266,683 the disclosure of which is incorporated by reference herein. The currently preferred matrices are xenogenic, allogenic or autogenic in nature. It is contemplated, however, that synthetic materials comprising polylactic acid, polyglycolic acid, polybutyric acid, derivatives and copolymers thereof to name but a few can also be used to generate suitable matrices. Preferred synthetic and naturally derived matrix materials, their preparation, methods for formulating them with the morphogenic proteins of the invention, and methods of administration are well known in the art and so are not discussed in detailed herein. See for example, U.S. Patent No. 5,266,683, the disclosure of which is herein incorporated by reference. It is further contemplated herein that binding to, adherence to or association with a matrix or the metal surface of a prosthetic device is an attribute that can be altered using the materials and methods disclosed herein. For example, devices comprising a matrix and osteoactive chimera of the instant invention having enhanced matrix-adherent properties can be used as a slow-release device. The skilled artisan will appreciate the variations and manipulations now possible in light of the teachings herein.

As will be appreciated by those skilled in the art, the concentration of the compounds described in a therapeutic composition will vary depending upon a number of factors, including the morphogenic effective amount to be administered, the chemical characteristics (e.g., hydrophobicity) of the compounds employed, and the route of administration. The preferred dosage of drug to be administered also is likely to depend on such variables as the type and

extent of a disease, tissue loss or defect, the overall health status of the particular patient, the relative biological efficacy of the compound selected, the formulation of the compound, the presence and types of excipients in the formulation, and the route of administration. In general terms, the therapeutic molecules of this invention may be provided to an individual where typical doses range from about 10 ng/kg to about 1 g/kg of body weight per day; with a preferred dose range being from about 0.1 mg/kg to 100 mg/kg of body weight.

II. SPECIFIC MODIFIED MORPHOGEN CONSTRUCTS

Generally, the modified morphogens of the present invention display chimeric attributes produced, for example, by domain-swapping regions of the finger 1 subdomain, which alters folding and/or alters activity of the morphogens. In certain preferred embodiments, the modified morphogens contain specific regions of the signature sequence of the finger 1 subdomain which impart altered receptor specificity to the morphogens, and alter stability, solubility as well as bioactivity and biospecificity.

According to the present invention, the attributes of native BMPs or other members of the TGF-β superfamily of proteins, including heterodimers and homodimers thereof, is altered by domain-swapping, for example, by exchanging defined regions of the F1 subdomain between proteins to alter one or more biological properties of a BMP or TGF-β superfamily member. As a result of this discovery, it is possible to design, TGF-β superfamily proteins that (1) are expressed recombinantly in prokaryotic or eukaryotic cells synthesized using polypeptide or nucleotide synthesizers; (2) have altered folding attributes; (3) have altered solubility under neutral pHs, including but not limited to physiologically compatible conditions; (4) have altered isoelectric points; (5) have altered stability; (6) have an altered tissue or receptor specificity; (7) have a re-designed, altered biological activity; and/or (8) have altered binding or adherence properties to solid surfaces, such as but not limited to, biocompatible matrices or metals. Thus, the present invention can provide mechanisms for designing quick-release, slow-release and/or timed-release formulations containing a preferred chimeric protein. Other advantages and features will be evident from the teachings below. Moreover, making use of the discoveries disclosed herein, chimeric proteins having altered surface-binding/surface-adherent properties can be designed and selected. Surfaces of particular significance include, but are not limited to,

solid surfaces which can be naturally-occurring such as bone; or porous particulate surfaces such as collagen or other biocompatible matrices; or the fabricated surfaces of prosthetic implants, including metals. As contemplated herein, virtually any surface can be assayed for differential binding of chimeric constructs. Thus, the present invention embraces a diversity of functional molecules having alterations in their surface-binding/surface-adherent properties, thereby rendering such constructs useful for altered *in vivo* applications, including slow-release, fast-release and/or timed-release formulations.

The skilled artisan will appreciate that mixing-and-matching any one or more the above-recited attributes provides specific opportunities to manipulate the uses of customized chimeric proteins (and DNAs encoding the same). For example, the attribute of altered stability can be exploited to manipulate the turnover of a chimeric protein *in vivo*. Moreover, in the case of chimeric proteins also having attributes such as altered re-folding and/or function, there is likely an interconnection between folding, function and stability. See, for example, Lipscomb et al., 7 Protein Sci. 765-73 (1998); and Nikolova et al., 95 Proc. Natl. Acad. Sci. USA 14675-80 (1998). For purposes of the present invention, stability alterations can be routinely monitored using well-known techniques of circular dichroism other indices of stability as a function of denaturant concentration or temperature. One can also use routine scanning calorimetry. Similarly, there is likely an interconnection between any of the foregoing attributes and the attribute of solubility. In the case of solubility, it is possible to manipulate this attribute so that a chimeric protein is either more or less soluble under physiologically-compatible and it consequently diffuses readily or remains localized, respectively, when administered *in vivo*.

In addition to the aforementioned uses of chimeric constructs with altered attributes, those with altered stability can also be used to practical advantage for shelf-life, storage and/or shipping considerations. Furthermore, on a related matter, altered stability can also directly affect dosage considerations thereby, for example, reducing the cost of treatment.

A particularly significant class of chimeric constructs are those having altered binding to solubilized carriers or excipients. By way of non-limiting example, an altered BMP having enhanced binding to a solubilized carrier such as hyaluronic acid permits the skilled artisan to administer an injectable formulation at a defect site without loss or dilution of the BMP by either diffusion or body fluids. Thus localization is maximized. The skilled artisan will appreciate the

variations made possible by the instant teachings. Similarly, another class of chimeric constructs having altered binding to body/tissue components can be exploited. By way of non-limiting example, an altered BMP having diminished binding to an in-situ inhibitor can be used to enhance repair of certain tissues *in vivo*. It is well known in the art, for example, that cartilage tissue is associated with certain proteins found in body fluids and/or within cartilage *per se* that can inhibit the activity of native BMPs. Chimeric constructs with altered binding properties, however, can overcome the effects of these in-situ inhibitors thereby enhancing repair, etc. The skilled artisan will appreciate the variations made possible by the instant teachings.

Figures 4A and 4B are schematic drawings of various biosynthetic constructs and their refolding and biological activity properties. In the figure, the thick solid line represents OP-1 (residues 330-431 of SEQ ID NO: 39), a thin solid line represents CDMP-2/GDF-6 (SEQ ID NO: 84), or GDF-5 (SEQ ID NO: 83), and a hatched line represents BMP-2 (SEQ ID NO: 49). Chimeric constructs are presented in Fig. 4A and individual sequence mutants are presented in Fig. 4B. In the figure, individual residue changes are indicated by the substituting residue in the figure. The residue numbering is as described above, counting from the first residue after the cysteine doublet in finger 2. For example, construct H2233 is OP-1(25R>E 26N>D 30R>E 35H>R). Similarly, construct H2447 has all the changes in H2233 and, in addition, has 1A>V 4Q>E 6N>S. Construct H2433 is OP-1(4Q>K 35H>R), and construct H2456 is OP-1(4Q>E 6N>S 25R>E 26N>D 30R>E 35H>R).

For each construct listed herein, refolding was measured using at least one of the parameters described herein (see, e.g., Example 4). All values recited were measured against a known good refolder, e.g., CDMP-2 or BMP-2, which yields at least 25% dimer upon refolding, without additional purification. Values are presented in Fig. 4 as follows: (+++) = >25%; (++) = 5-25%; (+) = 1-5%; (+/-) = < 1%, where native CDMP-2 and BMP-2 each have a +++ refolding value. Activity was measured in the osteosarcoma cell-based assay, which also is a measure of bone inducing activity. That is, in the assay, OP-1 and BMP-2 can induce alkaline phosphatase activity in this cell-based assay, but CDMP-2 and CDMP-1 do not. Yet, all four proteins are competent to induce the full cascade of morphogenic events resulting in endochondral bone formation in the rat subcutaneous assay (Example 8). In the cell-based assay of Example 5, activity measured is alkaline phosphatase activity, as determined by optical density at 405 nm.

The hOP-1 value shown is from mammalian cell –produced protein. "N/A" means "not-tested". Figure values are as follows: (+++) = $\geq$ 1.2; (++) 0.8-1.2; (+) = 0.4-0.8; (+/-) = < 0.4, where native hOP-1 has a +++ biological activity value.

As shown in Fig. 4B, replacing the C-terminal amino acid residue of OP-1 increases refolding of the protein. Chimeric constructs may also be produced that substitute heel or finger 1 sub-domains from good refolders. When not just finger-2 but also finger-1 was taken from CDMP, with just the heel of OP-1 in construct pH2388, the refolding was improved, showing that there is yet some context between the fingers. In construct pH2388, it was possible to insert just the heel of OP-1 into CDMP-2 and retain its good folding. This also indicated that the heel of OP-1 was not a problem with regard to refolding. The surprising fact that a foreign heel was easily accepted also showed that the three domains, which are demarcated by the cysteines, finger-1, heel, and finger-2 are structurally quite independent from each other.

Further, the finger 1 domain confers the specific activity of the morphogen constructs of the present invention. Two constructs, H2388 and H2389, which have the heel of OP-1 and either both fingers from CDMP-2 or only finger-2 of CDMP-2 had been successfully refolded and were tested in the ROS assay for alkaline phosphatase induction and in the rat implant assay for bone formation. No alkaine phosphatase induction was found in H2388 or wt CDMP-2. However, activity was induced by H2389. This is illustrated in Fig. 4A. Also, OP-1 with the finger-2 domain of BMP-2, H2410, folds well and has good activity. Thus ROS activity is indiscriminate with respect to which finger 2 (OP-1, BMP-2, or CDMP-2) is present, as long as it is intact. But in order to have ROS activity, a finger-1 domain from a OP-1 or BMP-2, or other ROS-active BMP is most preferably present in certain modified morphogens.

Practice of the invention will be still more fully understood from the following examples, which are presented herein for illustration only and should not be construed as limiting the invention in any way.

III. EXAMPLES

EXAMPLE 1. Synthesis of an Exemplary Chimeric Protein of the Present Invention: a BMP Mutant

Figure 3 shows the nucleotide and corresponding amino acid sequence for the OP-1 C-terminal seven cysteine domain. Knowing these sequences permits identification of useful restriction sites for engineering in mutations by, for example, cassette mutagenesis or the well-known method of Kunkel (mutagenesis by primer extension using m13-derived single-stranded templates) or by the well-known PCR methods, including overlap extension. An exemplary mutant of OP-1 is H2460, with 4 amino acid changes in the finger 2 sub-domain and an amino acid change in the last C-terminal amino acid, constructed as described below. It is understood by the skilled artisan that the mutagenesis protocol described is exemplary only, and that other means for creating the constructs of the invention are well-known and well described in the art.

Four amino acid changes were introduced into the OP-1 finger 2 sub-domain sequence by means of standard polymerase chain reactions using overlap extension technique, resulting in OP-1 mutant H2460. The four changes in the finger 2 region were N6>S, R25>E, N26>D and R30>E. This mutant also contained a further change, H35>R, of the C-terminal residue. The template for these reactions was the mature domain of a wild type OP-1 cDNA clone, which had been inserted into an E.coli expression vector engineered with an ATG start codon at the beginning of the mature region. The ATG had been introduced by PCR using as a forward primer a synthetic oligonucleotide of the following sequence: ATG TCC ACG GGG AGC AAA CAG (SEQ ID NO: 36), encoding M S T G S K Q (SEQ ID NO: 37). The PCR reaction was done in combination with an appropriate back-primer complementary to the 3' coding region of the cDNA.

In order to construct the finger 2 mutant H2460, a PCR fragment encoding the modified finger-2 was made in a standard PCR reaction, using a commercially available PCR kit and following the manufacturer's instructions using as primers synthetic oligonucleotides.

To obtain the N6>S change, a forward primer (primer #1) of the sequence GCG CCC ACG CAG CTC AGC GCT ATC TCC GTC CTC (SEQ ID NO: 70) was used, encoding the amino acid sequence: A P T Q L S A I S V L (SEQ ID NO: 71).

For the changes near the C-terminus, a back-primer, 43 nucleotides long, (primer #2) was used which introduced the R25>E and N26>D and R30>E and C-terminal H35>R changes. This primer #2 had the sequence: CTA TCT GCA GCC ACA AGC TTC GAC CAC CAT GTC TTC GTA TTT C (SEQ ID NO: 72) which is the complement of the coding sequence, G AAA TAC

GAA GAC ATG GTG GTC GAA GCT TGT GGC TGC AGA TAG (SEQ ID NO: 73) encoding the amino acids: K Y E D M V V E A C G C R stop (SEQ ID NO: 74).

The fragment with finger 2 and C-terminus mutations was then combined with another PCR fragment encoding the upstream part of mature OP-1, with N-terminus, finger-1 and heel sub-domains. The latter PCR fragment, encoding the N-terminus, finger 1 and heel sub-domains was constructed again using an OP-1 expression vector for E.coli as template. The vector contained an OP-1 cDNA fragment, encoding the mature OP-1 protein attached to a T7 promoter and ribosome binding site for expression under control of either a T7 promoter in an appropriate host or under control of a trp promoter. In this T7 expression vector, Pet 3d (Novagen Inc., Madison WI) the sequence between the T7 promoter, at the XbaI site, and the ATG codon of mature OP-1 is as follows: TCTAGAATAATTTTGTTTAACCTTTAAGAAGGAGATATACG ATG (SEQ ID NO: 75).

This second PCR reaction was primed with a forward primer (primer #3) TAA TAC GAC TCA CTA TAG G (SEQ ID NO: 76) which primes in the T7 promoter region and a back-primer (primer #4) that overlaps with primer #1 and has the nucleotide sequence GCT GAG CTG CGT GGG CGC (SEQ ID NO: 77), which is the complement of the coding sequence GCG CCC ACG CAG CTC AGC (SEQ ID NO: 78), encoding A P T Q L S (SEQ ID NO:79).

In a third PCR reaction, the actual overlap extension reaction, portions of the above two PCR fragments were combined and amplified by PCR, resulting in a single fragment containing the complete mature OP-1 region. For this reaction, primer #3 was used as forward primer and a new primer (primer #5) was used as a back-primer with the following sequence GG ATC CTA TCT GCA GCC ACA AGC (SEQ ID NO: 80), which is the complement to coding sequence GCT TGT GGC TGC AGA TAG GAT CC (SEQ ID NO: 81), encoding A C G C R stop (SEQ ID NO: 82). This primer also adds a convenient 3' BamHI site for of inserting the gene into the expression vector.

The resulting fragment bearing the complete mutant gene, resulting from the overlap extension PCR, was cloned into a commercial cloning vector designed for cloning of PCR fragments, such as pCR2.1-topo-TA (Invitrogen Inc., Carlsbad CA). The cloned PCR fragment was recovered by restriction digest with XbaI and BamHI and inserted into the XbaI and BamHI

sites of a commercially available T7 expression vector such as Pet3d (Novagen Inc., Madison WI).

EXAMPLE 2. *E. coli* Expression of an Exemplary Chimeric Protein of the Present Invention: Expression of a BMP Mutant

Transformed cells were grown in standard SPYE 2YT media, 1:1 ratio, (see, Sambrook et al., for example) at 37°C, under standard culturing conditions. Heterologous protein overexpression typically produced inclusion bodies within 8-48 hours. Inclusion bodies were isolated and solubilized as follows. One liter of culture fluid was centrifuged to collect the cells. The cells in the resulting pellet then were resuspended in 60 ml 25 mM Tris, 10 mM EDTA, pH 8.0 (TE Buffer) + 100 μg/ml lysozyme and incubated at 37°C for 2 hours. The cell suspension was then chilled on ice and sonicated to lyse the cells. Cell lysis was ascertained by microscopic examination. The volume of the lysate was adjusted to approximately 300 ml with TE Buffer, then centrifuged to obtain an inclusion body pellet. The pellet was washed by 2-4 successive resuspensions in TE Buffer and centrifugation. The washed inclusion body pellet was solubilized by denaturation and reduction in 40 ml 100 mM Tris, 10 mM EDTA, 6M GuHCl (guanidinium hydrochloride), 250 mM DTT, pH 8.8. Proteins then were pre-purified using a standard, commercially available C2 or C8 cartridge (SPICE cartridges, 400 mg, Analtech, Inc.). Protein solutions were acidified with 2% TFA (trifluoroacetic acid), applied to the cartridge, washed with 0.1% TFA/10% acetonitrile, and eluted with 0.1%TFA/70% acetonitrile. The eluted material then was dried down or diluted and fractionated by C4 RP-HPLC.

EXAMPLE 3. Refolding of an Exemplary Chimeric Protein of the Present Invention: a Mutant BMP Dimer

Proteins prepared as described above were dried down prior to refolding, or diluted directly into refolding buffer. The preferred refolding buffer used was: 100 mM Tris, 10 mM EDTA, 1 M NaCl, 2% CHAPS, 5 mM GSH (reduced glutathione), 2.5 mM GSSG (oxidized glutathione), pH 8.5. Refoldings (12.5-200 μg protein/ml) were carried out at 4°C for 24-90 hours, typically 36-48 hours, although longer than this (up to weeks) are expected to provide good refolding in some mutants, followed by dialysis against 0.1% TFA, then 0.01% TFA, 50% ethanol. Aliquots of the dialyzed material then was dried down in preparation for the various

assays.

EXAMPLE 4. Purification and Testing of a Refolded Exemplary Chimeric Dimer of the Present Invention: a Mutant BMP Dimer

4A. SDS-PAGE, RP HPLC –

Samples were dried down and resuspended in Laemmli gel sample buffer and then electrophoresed in a 15% SDS-polyacrylamide gel. All assays included molecular weight standards and/or purified mammalian cell produced OP-1 for comparison. Analysis of OP-1 dimers was performed in the absence of added reducing agents, while OP-1 monomers were produced by the addition of 100 mM DTT to the gel samples. Folded dimer has an apparent molecular weight in the range of about 30-36 kDa, while monomeric species have an apparent molecular weight of about 14-16 kDa.

Alternatively, samples were chromatographed on a commercially available RP-HPLC, as follows. Samples were dried down and resuspended in 0.1% TFA/30% acetonitrile. The protein then was applied to a C18 column in 0.1% TFA, 30% acetonitrile and fractionated using a 30-60% acetonitrile gradient in 0.1% TFA. Properly folded dimers elute as a discrete peak at 45-50% acetonitrile; monomers elute at 50-60% acetonitrile.

4B. Hydroxyapatite Chromatography –

Samples were loaded onto hydroxyapatite column in 10mM phosphate, 6 M urea, pH 7.0 (Column Buffer). Unbound material was removed by washing with column buffer, followed by elution of monomer with Column Buffer + 100 mM NaCl. Dimers were eluted with Column Buffer + 250 mM NaCl. .

4C. Trypsin Digest –

Tryptic digests were performed in a digestion buffer of 50 mM Tris, 4 M urea, 100 mM NaCl, 0.3% Tween 80, 20 mM methylamine, pH 8.0. The ratio of enzyme to substrate was 1:50 (weight to weight). After incubation at 37°C for 16 hours, 15 µl of digestion mixture was combined with 5 µl 4X gel sample buffer without DTT and analyzed by SDS-PAGE. Purified mammalian OP-1 and undigested BMP dimer were included for comparison. Under these conditions, properly folded dimers are cleaved to produce a species with slightly faster migration

than uncleaved standards, while monomers and mis-folded dimers are completely digested and do not appear as bands in the stained gel.

EXAMPLE 5. Testing the Altered Attributes of a Chimeric Protein of the Present Invention: *In vitro* Cell-Based Bioassay of Osteogenic Activity

This example demonstrates the bioactivity of certain osteogenic chimeras proteins of the invention. Native osteogenic proteins having high specific bone forming activity can induce alkaline phosphatase activity in rat osteoblasts, including rat osteosarcoma cells and rat calveria cells. In the assay rat osteosarcoma or calveria cells were plated onto a multi-well plate (e.g., a 48 well plate) at a concentration of 50,000 osteoblasts per well, in αMEM (modified Eagle's medium, Gibco, Inc. Long Island) containing 10% FBS (fetal bovine serum), L-glutamine and penicillin/streptomycin. The cells were incubated for 24 hours at 37°C, at which time the growth medium was replaced with α MEM containing 1% FBS and the cells incubated for an additional 24 hours so that cells were in serum-deprived growth medium at the time of the experiment. Cultured cells then were divided into three groups: (1) wells receiving various concentrations of biosynthetic osteogenic protein; (2) a positive control, such as mammalian expressed hOP-1; and a negative control (no protein or TGF-β). The protein concentrations tested were in the range of 50-500 ng/ml. Cells were incubated for 72 hours. After the incubation period the cell layer was extracted with 0.5 ml of 1% TritonX-100. The resultant cell extract was centrifuged, 100 µl of the extract was added to 90 µl of PNPP (paranitrosophenylphosphate)/glycerine mixture and incubated for 30 minutes in a 37°C water bath and the reaction stopped with 100 µl NaOH. The samples then were run through a plate reader (e.g., Dynatech MR700) and absorbance measured at 400 nm, using p-nitrophenol as a standard, to determine the presence and amount of alkaline phosphatase activity. Protein concentrations were determined by standard means, e.g., the Biorad method, UV scan or RP HPLC area at 214 nm. Alkaline phosphatase activity was calculated in units/µg protein, where 1 unit equals 1 nmol p-nitrophenol liberated/30 minutes at 37°C.

Native hOP-1 and BMP2 generate approximately 1.0-1.4 units at between 100-200 ng/ml.

EXAMPLE 6. Testing the Altered Attributes of a Chimeric Protein of the Present Invention: *In vitro* Cell-Based Bioassay of CDMP Activity

This example demonstrates the bioactivity of certain osteogenic chimeras of the invention. Native CDMPs fail to induce alkaline phosphatase activity in rat osteosarcoma cells as used in Example 5, but they do induce alkaline phosphatase activity in the mouse teratocarcinoma cell line ATDC-5, a chondroprogenitor cell line (Atsumi, et al., 1990, *Cell Differentiation and Development 30*: 109). Re-folded chimeric constructs that are negative in the rat osteocarcinoma cell assay but positive in the ATDC-5 assay are described as having acquired CDMP-like activity. In the ATDC-5 assay, cells were plated at density of 4×10^4 in serum-free basal medium (BM: Ham's F-12/DMEM [1:1] with ITS™ + culture supplement [Collaborative Biomedical Products, Bedford, MA], alpha-ketoglutarate (1×10^{-4} M), ceruloplasmin (0.25 U/ml), cholesterol (5 μg/ml), phosphatidylethanolamine (2 μg/ml), alpha-tocopherol acid succinate (9×10^{-7} M), reduced glutathione (10 μg/ml), taurine (1.25 μg/ml), triiodothyronin (1.6×10^{-9} M), hydrocortisone (1×10^{-9} M), parathyroid hormone (5×10^{-10} M), β-glycerophosphate (10 mM), and L-ascorbic acid 2-sulphate (50 μg/ml)). CDMP or chimeric protein (0 – 300 ng/ml) was added the next day and the culture medium, including CDMP or chimeric protein, replaced every other day. Alkaline phosphatase activity was determined in sonicated cell homogenates after 4, 6 and/or 12 days of treatment. After extensive washing with PBS, cell layers were sonicated in 500 μl of PBS containing 0.05% Triton-X100. 50-100μl aliquots were assayed for enzyme activity in assay buffer (0.1M sodium barbital buffer, pH 9.3) and p-nitrophenyl phosphate as substrate. Absorbance was measured at 400 nm, and activity normalized to protein content measured by Bradford protein assay (bovine serum albumin standard).

Native CDMP-1 and CDMP-2 generated approximately 2-3 units of activity at day 10 at 100 ng/ml. Native OP-1 generated approximately 6-7 units of activity at day 10 at 100 ng/ml.

EXAMPLE 7. Testing the Altered Attributes of a Chimeric Protein of the Present Invention: *In vitro* Cell-Based Bioassay of TGF-ß-like Activity

This example demonstrates the bioactivity of certain TGF-β-based chimeric proteins of the invention. Native TGF-β proteins can inhibit epithelial cell proliferation. Numerous cell inhibition assays are well described in the art. See, for example, Brown, et. al. (1987) J.

Immunol. 139:2977, describing a colorimetric assay using human melanoma A375 fibroblast cells, and described herein below. Another assay uses epithelial cells, e.g., mink lung epithelial cells, and proliferative effects are determined by ^{3}H-thymidine uptake.

Briefly, in the assay the TGF-β biosynthetic construct is serially diluted in a multi-well tissue plate containing RPMI-1640 medium (Gibco) and 5% fetal calf serum. Control wells receive medium only. Melanoma cells then are added to the well (1.5×10^{4}). The plates then are incubated at 37°C for about 72 hours in 5% CO_2, and the cell monolayers washed once, fixed and stained with crystalviolet for 15 minutes. Unbound stain is washed out and the stained cells then lysed with 33% acetic acid to release the stain (confined to the cell nuclei), and the OD measured at 590 nm with a standard, commercially available photometer to calculate the activity of the test molecules. The intensity of staining in each well is directly related to the number of nuclei. Accordingly, active TGF-β molecules are expected to stain lighter than inactive compounds or the negative control well.

In another assay, mink lung cells are used. These cells grow and proliferate under standard culturing conditions, but are arrested following exposure to TGF-ß, as determined by ^{3}H-thymidine uptake using culture cells from a mink lung epithelial cell line (ATCC No. CCL 64, Rockville, MD). Briefly cells are grown to confluency with in EMEM, supplemented with 10% FBS, 200 units/ml penicillin, and 200 µg/ml streptomycin. These cells are cultured to a cell density of about 200,000 cells per well. At confluency the media is replaced with 0.5 ml of EMEM containing 1%FBS and penicillin/streptomycin and the culture incubated for 24 hours at 37°C. Candidate proteins then are added to each well and the cells incubated for 18 hours at 37° C. After incubation, 1.0 µCi of ^{3}H-thymidine in 10 µl was added to each well, and the cells incubated for four hours at 37°C. The media then is removed from each well and the cells washed once with ice-cold phosphate buffered saline and DNA precipitated by adding 0.5 ml of 10% TCA to each well and incubated at room temperature for 15 minutes. The cells are washed three times with ice-cold distilled water, lysed with 0.5 ml 0.4 M NaOH, and the lysate from each well then transferred to a scintillation vial and the radioactivity recorded using a scintillation counter (Smith-Kline Beckman). Biologically active molecules will inhibit cell proliferation resulting in less thymidine uptake and fewer counts as compared to inactive proteins and/or the negative control well (no added growth factor).

EXAMPLE 8. Testing the Altered Attributes of a Chimeric Protein of the Present Invention: *In vivo* Bioassay of Osteogenic Activity (Endochondral Bone Formation and Related Properties)

The art-recognized bioassay for bone induction as described by Sampath and Reddi (Proc. Natl. Acad. Sci. USA (1983) 80:6591-6595) and US Pat. Nos. 4,968,590, 5,266,683, the disclosures of which is herein incorporated by reference, can be used to establish the efficacy of a given device or formulation. Briefly, the assay consists of depositing test samples in subcutaneous sites in recipient rats under ether anesthesia. A vertical incision (1 cm) is made under sterile conditions in the skin over the thoracic region, and a pocket is prepared by blunt dissection. In certain, to the desired amount of osteogenic protein (10 ng – 10 μg) approximately 25 mg of matrix material, using standard procedures such as lyophilization, is added and the test sample is implanted deep into the pocket and the incision is closed with a metallic skin clip. The heterotropic site allows for the study of bone induction without the possible ambiguities resulting from the use of orthotopic sites. The implants also can be provided intramuscularly which places the devices in closer contact with accessable progenitor cells. Typically intramuscular implants are made in the skeletal muscle of both legs.

The sequential cellular reactions occurring at the heterotropic site are complex. The multistep cascade of endochondral bone formation includes: binding of fibrin and fibronectin to implanted matrix, chemotaxis of cells, proliferation of fibroblasts, differentiation into chondroblasts, cartilage formation, vascular invasion, bone formation, remodeling, and bone marrow differentiation.

Successful implants exhibit a controlled progression through the stages of protein-induced endochondral bone development including: (1) transient infiltration by polymorphonuclear leukocytes on day one; (2) mesenchymal cell migration and proliferation on days two and three; (3) chondrocyte appearance on days five and six; (4) cartilage matrix formation on day seven; (5) cartilage calcification on day eight; (6) vascular invasion, appearance of osteoblasts, and formation of new bone on days nine and ten; (7) appearance of osteoblastic and bone remodeling on days twelve to eighteen; and (8) hematopoietic bone marrow differentiation in the ossicle on day twenty-one.

Histological sectioning and staining is preferred to determine the extent of osteogenesis in the implants. Staining with toluidine blue or hemotoxylin/eosin clearly demonstrates the ultimate development of endochondral bone. Twelve day bioassays are sufficient to determine whether bone inducing activity is associated with the test sample.

Additionally, alkaline phosphatase activity and/or total calcium content can be used as biochemical markers for osteogenesis. The alkaline phosphatase enzyme activity can be determined spectrophotometrically after homogenization of the excised test material. The activity peaks at 9-10 days *in vivo* and thereafter slowly declines. Samples showing no bone development by histology should have no alkaline phosphatase activity under these assay conditions. The assay is useful for quantitation and obtaining an estimate of bone formation very quickly after the test samples are removed from the rat. The results as measured by alkaline phosphatase activity level and histological evaluation can be represented as "bone forming units". One bone forming unit represents the amount of protein that is needed for half maximal bone forming activity on day 12. Additionally, dose curves can be constructed for bone inducing activity *in vivo* at each step of a purification scheme by assaying various concentrations of protein. Accordingly, the skilled artisan can construct representative dose curves using only routine experimentation.

Total calcium content can be determined after homogenization in, for example, cold 0.15M NaCl, 3 mM $NaHCO_3$, pH 9.0, and measuring the calcium content of the acid soluble fraction of sediment.

EXAMPLE 9: Exemplary Modified Morphogen Constructs Having Altered Attributes: Domain Swapping

Figure 4 is a schematic drawing of various biosynthetic constructs and their refolding and biological attributes. The thick solid line represents OP-1, a thin solid line represents CDMP-2/GDF-6 or GDF-5, and a hatched line represents BMP-2. Chimeric constructs are presented in Fig. 4A and individual sequence mutants are presented in Fig. 4B. Individual residue changes are indicated by the substituting residue in the figure. The residue numbering is as described above, counting from the first residue after the cysteine doublet in finger 2. For example, construct H2233 is OP-1(25R>E 26N>D 30R>E 35H>R). Similarly, construct H2447 has all the

changes in H2233 and, in addition, has 1A>V 4Q>E 6N>S. Construct H2433 is OP-1(4Q>K 35H>R), and construct H2456 is OP-1(4Q>E 6N>S 25R>E 26N>D 30R>E 35H>R).

As taught herein, the present invention provides the skilled artisan with the know-how to craft customized chimeric proteins and DNAs encoding the same. Further taught and exemplified herein are the means to design chimeric proteins having certain desired attribute(s) making them suitable for specific *in vivo* applications (see at least Sections I.B.,II., and III. Examples 1-4, 8 and 11 for exemplary embodiments of the foregoing chimeric proteins). For example, chimeric proteins having altered solubility attributes can be used *in vivo* to manipulate morphogenic effective amounts provided to a recipient. That is, increased solubility can result in increased availability; diminished solubility can result in decreased availability. Thus, such systemically administered chimeric proteins can be immediately available/have immediate morphogenic effects, whereas locally administered chimeric proteins can be available more slowly/have prolonged morphogenic effects. The skilled artisan will appreciate when increased *versus* diminished solubility attributes are preferred given the facts and circumstances at hand. Optimization of such parameters requires routine experimentation and ordinary skill.

Similarly, chimeric proteins having altered stability attributes can be used *in vivo* to manipulate morphogenic effective amounts provided to a recipient. That is, increased stability can result in increased half-life because turnover *in vivo* is less; diminished stability can result in decreased half-life and availability because turnover *in vivo* is more. Thus, such systemically administered chimeric proteins can either be immediately available/have immediate morphogenic effects achieving a bolus-type dosage or can be available *in vivo* for prolonged periods/have prolonged morphogenic effects achieving a sustained release type dosage. The skilled artisan will appreciate when increased *versus* diminished stability attributes are preferred given the facts and circumstances at hand. Optimization of such parameters requires routine experimentation and ordinary skill.

Additionally, chimeric proteins having a combination of altered attributes, such as but not limited to solubility and stability attributes, can be used *in vivo* to manipulate morphogenic effective amounts provided to a recipient. That is, by designing a chimeric protein with a combination of specific altered attributes, morphogenic effective amounts can be administered in

a timed-release fashion; dosages can be regulated both in terms of amount and duration; treatment regimens can be initiated at low doses systemically or locally followed by a transition to high doses, or *vice versa*; to name but a few paradigms. The skilled artisan will appreciate when low *versus* high morphogenic effective amounts are suitable under the facts and circumstances at hand. Optimization of such parameters requires routine experimentation and ordinary skill.

Furthermore, chimeric proteins having one or more altered attributes are useful to overcome inherent deficiencies in development. Chimeric proteins having one or more altered attributes can be designed to circumvent an inherent defect in a host's native morphogenic signaling system. As a non-limiting example, a chimeric protein of the present invention can be used to bypass a defect in a native receptor in a target tissue, a defect in an intracellular signaling pathway, and/or a defect in other events which are reliant on the attributes of a subdomain(s) associated with recognition of a moiety *per se* as opposed to the attributes associated with function/biological activity which are embodied in a different subdomain(s). The skilled artisan will appreciate when such chimeric proteins are suitable given the facts and circumstances at hand. Optimization requires routine experimentation and ordinary skill.

Example 10. Determination of Binding Activity of an OP-1 Based Chimeric Construct

Cells expressing an OP-1 receptor on their cell surface are plated into 35 mm dishes and incubated for 48 hours in DMEM (Dulbecco's modified Eagle medium) plus 10% fetal calf serum. Purified OP-1, or an OP-1-analog is iodinated with $Na^{125}I$ by chloramine T oxidation, preferably having a specific activity of about 50-100 mCi/mg, by following essentially the protocol of Frolik et al. (1984) J. Biol. Chem. 595: 10995-11000. Labeled OP-1, or OP-1-analog, then is purified using a standard procedure, e.g., by chromatographic separation. The plated cells then are washed twice with physiologically buffered saline in the presence of 0.1% BSA, and incubated at 22°C in the presence of BSA, buffer and labeled OP-1 (1 ng) and various concentrations (e.g., 0-10mg/ml) of unlabelled competitor, e.g., unlabelled OP-1 or the OP-1 based construct. Following binding, the cells are washed three times with cold buffer, solubilized in 0.5 ml of 0.5 N NaOH, removed from the dish, and radioactivity determined by

gamma or scintillation counter. Data then are expressed as percent inhibition, where 100% inhibition of specific binding is the difference between binding in the absence of competitor and binding in the presence of a 100-fold molar excess of unlabelled competitor. Binding parameters preferably are determined using a computer program such as LIGAND (Munsun et al. (1980) Anal. Biochem. 107: 220-259). Upon performance of the assay, it is contemplated that the OP-1 based construct may have specific binding activity for the OP-1 receptor. Upon confirmation of binding activity, the construct can be tested subsequently for other indicia of biological activity.

Example 11. Other Indicia of Biological Activities of an OP-1 Based Chimeric Construct.

The biological activity of the resulting OP-1 based chimeric construct can be determined using any of the using standard *in vivo* and *in vitro* assays, typically used for assessing native OP-1 activity. A variety of additional exemplary assays are set forth in detail below.

A. Progenitor Cell Stimulation.

The following example is designed to demonstrate the ability of OP-1 based constructs to stimulate the proliferation of mesenchymal progenitor cells. Useful naive stem cells include pluripotential stem cells, which may be isolated from bone marrow or umbilical cord blood using conventional methodologies, (see, for example, Faradji et al. (1988) Vox Sang. 55 (3): 133-138 or Broxmeyer et al. (1989) Proc. Natl. Acad. Sci. USA. 86: 3828-3832), as well as naive stem cells obtained from blood. Alternatively, embryonic cells (e.g., from a cultured mesodermal cell line) may be used.

Another method for obtaining progenitor cells and for determining the ability of OP-1 based constructs to stimulate cell proliferation is to capture progenitor cells from an *in vivo* source. For example, a biocompatible matrix material able to allow the influx of migratory progenitor cells may be implanted at an *in vivo* site long enough to allow the influx of migratory progenitor cells. For example, a bone-derived, guanidine-extracted matrix, formulated as disclosed for example in Sampath et al. (1983) Proc. Natl. Acad. Sci. USA 80: 6591-6595, or U.S. Patent No. 4,975,526, may be implanted into a rat at a subcutaneous site, essentially following the method of Sampath et al. After three days the implant is removed, and the progenitor cells associated with the matrix dispersed and cultured.

Progenitor cells, however obtained, then are incubated *in vitro* with the OP-1 based construct under standard cell culture conditions, such as those described hereinbelow. In the absence of external stimuli, the progenitor cells do not, or only minimally, proliferate on their own in culture. It is contemplated, however, that progenitor cells cultured in the presence of a biologically active OP-1 construct, like OP-1, will proliferate. Cell growth can be determined visually or spectrophotometrically using standard methods well known in the art.

B. Morphogen-Induced Cell Differentiation.

A variety of assays may be used to determine OP-1 based cellular differentiation.

1. Embryonic Mesenchyme Differentiation

As with natural OP-1, it is contemplated that OP-1 based constructs can be utilized to induce cell differentiation. The ability to induce cell differentiation can be demonstrated by culturing early mesenchymal cells in the presence of the chimeric protein and then studying the histology of the cultured cells by staining with toluidine blue using standard cell culturing and cell staining methodologies well described in the art. For example, it is known that rat mesenchymal cells destined to become mandibular bone, when separated from the overlying epithelial cells at stage 11 and cultured *in vitro* under standard tissue culture conditions, e.g., in a chemically defined, serum-free medium, containing for example, 67% DMEM (Dulbecco's modified Eagle's medium), 22% F-12 medium, 10mM Hepes pH 7, 2mM glutamine, 50 mg/ml transferring, 25 mg/ml insulin, trace elements, 2mg/ml bovine serum albumin coupled to oleic acid, with HAT (0.1 mM hypoxanthine, 10mM aminopterin, 12 mM thymidine), will not continue to differentiate. However, if these same cells are left in contact with the overlying endoderm for an additional day, at which time they become stage 12 cells, they will continue to differentiate on their own *in vitro* to form chondrocytes. Further differentiation into osteoblasts and, ultimately, mandibular bone, requires an appropriate local environment, e.g., a vascularized environment.

It is anticipated that, as with natural OP-1, stage 11 mesenchymal cells, cultured *in vitro* in the presence of an OP-1 chimera, e.g., 10-100 ng/ml, will continue to differentiate *in vitro* to form chondrocytes just as they continue to differentiate *in vitro* if they are cultured with the cell

products harvested from the overlying endodermal cells. This experiment may be performed with different mesenchymal cells to demonstrate cell differentiation capability.

2. Alkaline Phosphatase Induction of Osteoblasts.

As another example of morphogen-induced cell differentiation, the ability of OP-1 chimerics to induce osteoblast differentiation may be demonstrated *in vitro* using primary osteoblast cultures, or osteoblast-like cell lines, and assaying for a variety of bone cell markers that are specific markers for the differentiated osteoblast phenotype, e.g., alkaline phosphatase activity, parathyroid hormone-mediated cyclic AMP (cAMP) production, osteocalcin synthesis, and enhanced mineralization rates.

Cultured osteoblasts in serum-free medium are incubated with, a range of OP-1 chimera concentrations, for example, 0.1, 1.0, 10.0, 40.0 or 80.0 ng OP-1 chimera/ml medium; or with a similar concentration range of natural OP-1 or TGF-β. After a 72 hour incubation the cell layer is extracted with 0.5 ml of 1% Triton X-100. The resultant cell extract, is centrifuged, and 100 µl of the extract is added to 90 µl of paranitroso-phenylphosphate (PNPP)/glycerine mixture and incubated for 30 minutes in a 37°C water bath and the reaction stopped with 100 µl 0.2N NaOH. The samples then are run through a plate reader (e.g., Dynatech MR700 plate reader, and absorbance measured at 400 nm, using p-nitrophenol as a standard) to determine the presence and amount of alkaline phosphate activity. Protein concentrations are determined by the BioRad method. Alkaline phosphatase activity is calculated in units/µg protein, where 1 unit=1 nmol p-nitrophenol liberated/30 minutes at 37°C.

The OP-1 chimera, like natural OP-1 alone, stimulates the production of alkaline phosphatase in osteoblasts thereby promoting the growth and expression of the osteoblast differentiated phenotype.

The long term effect of OP-1 chimera on the production of alkaline phosphatase by rat osteoblasts also can be demonstrated as follows.

Rat osteoblasts are prepared and cultured in multi-well plates as described above. In this example six sets of 24 well plates are plated with 50,000 rat osteoblasts per well. The wells in each plate, prepared as described above, then are divided into three groups: (1) those which receive, for example, 1 ng of OP-1 chimera per ml of medium; (2) those which receive 40 ng of

OP-1 chimera per ml of medium; and (3) those which received 80 ng of OP-1 chimera per ml of medium. Each plate then is incubated for different lengths of time: 0 hours (control time), 24 hours, 48 hours, 96 hours, 120 hours and 144 hours. After each incubation period, the cell layer is extracted with 0.5 ml of 1% Triton X-100. The resultant cell extract is centrifuged, and alkaline phosphatase activity determined using paranitroso-phenylphosphate (PNPP), as above. The OP-1 chimera, like natural OP-1, will stimulate the production of alkaline phosphatase in osteoblasts in dose-dependent manner so that increasing doses of OP-1 chimera will further increase the level of alkaline phosphatase production. Moreover, it is contemplated that the OP-1 chimera-stimulated elevated levels of alkaline phosphatase in the treated osteoblasts will last for an extended period of time.

Substitute Sheet (Rule 26)

3. Induction of PTH-Mediated cAMP.

This experiment is designed to test the effect of OP-1 chimeras on parathyroid hormone-mediated cAMP production in rat osteoblasts *in vitro*. Briefly, rat osteoblasts are prepared and cultured in a multiwell plate as described above. The cultured cells then are divided into four groups: (1) wells which receive, for example, 1.0, 10.0 and 40.0 ng OP-1 mutant/ml medium); (2) wells which receive for example, natural OP-1, at similar concentration ranges; (3) wells which receive for example, TGF-β, at similar concentration ranges; and (4) a control group which receives no growth factors. The plate then is incubated for another 72 hours. At the end of the 72 hours the cells are treated with medium containing 0.5% bovine serum albumin (BSA) and 1mM 3-isobutyl-1-methylxanthine for 20 minutes followed by the addition into half of the wells of human recombinant parathyroid hormone (hPTH, Sigma, St. Louis) at a concentration of 200 ng/ml for 10 minutes. The cell layer then is extracted from each well with 0.5 ml of 1% Triton X-100. The cAMP levels then are determined using a radioimmunoassay kit (e.g., Amersham, Arlington Heights, Illinois). The OP-1 chimera alone, like OP-1, will stimulate an increase in the PTH-mediated cAMP response, thereby promoting the growth and expression of the osteoblast differentiated phenotype.

4. Induction of Osteocalcin Production.

Osteocalcin is a bone-specific protein synthesized by osteoblasts which plays an integral role in the rate of bone mineralization *in vivo*. Circulating levels of osteocalcin in serum are used as a marker for osteoblast activity and bone formation *in vivo*. Induction of osteocalcin synthesis in osteoblast-enriched cultures can be used to demonstrate OP-1 chimera efficacy *in vitro*.

Rat osteoblasts are prepared and cultured in a multi-well plate as above. In this experiment the medium is supplemented with 10%FBS, and on day 2, cells are fed with fresh medium supplemented with fresh 10 mM β-glycerophosphate (Sigma, Inc.). Beginning on day 5 and twice weekly thereafter, cells are fed with a complete mineralization medium containing all of the above components plus fresh L(+)-ascorbate, at a final concentration of 50mg/ml medium. OP-1 chimera then is added to the wells directly, e.g., in 50% acetonitrile (or 50% ethanol) containing 0.1% trifluoroacetic acid (TFA), at no more than 5ml chimera/ml medium. Control wells receive solvent vehicle only. The cells then are re-fed and the conditioned medium sample diluted 1:1 in standard radioimmunoassay buffer containing standard protease inhibitors and

stored at -20°C until assayed for osteocalcin. Osteocalcin synthesis is measured by standard radioimmunoassay using a commercially available osteocalcin-specific antibody.

Mineralization is determined on long term cultures (13 day) using a modified von Kossa staining technique on fixed cell layers: cells are fixed in fresh 4% paraformaldehyde at 23°C for 10 min, following rinsing cold 0.9% NaCl. Fixed cells then are stained for endogenous alkaline phosphatase at pH 9.5 for 10 min, using a commercially available kit (Sigma, Inc.). Purple stained cells then are dehydrated with methanol and air dried. After 30 min incubation in 3% $AgNO_3$ in the dark, H_2O-rinsed samples are exposed for 30 sec to 254 nm UV light to develop the black silver-stained phosphate nodules. Individual mineralized foci (at least 20 mm in size) are counted under a dissecting microscope and expressed as nodules/culture.

The OP-1 chimera, like natural OP-1, will stimulate osteocalcin synthesis in osteoblast cultures. Furthermore, the increased osteocalcin synthesis in response to OP-1 chimera will be in a dose dependent manner thereby showing a significant increase over the basal level after 13 days of incubation. The enhanced osteocalcin synthesis also may be confirmed by detecting the elevated osteocalcin mRNA message (20-fold increase) using a rat osteocalcin-specific probe. In addition, the increase in osteoclacin synthesis correlates with increased mineralization in long term osteoblast cultures as determined by the appearance of mineral nodules. It is contemplated also that OP-1 chimera, like natural OP-1, will increase significantly the initial mineralization rate as compared to untreated cultures.

5. Morphogen-Induced CAM Expression.

Native members of the vg/dpp subgroup induce CAM expression, particularly N-CAM expression, as part of their induction of morphogenesis (see copending USSN 922,813). CAMs are morphoregulatory molecules identified in all tissues as an essential step in tissue development. N-CAMs, which comprise at least 3 isoforms (N-CAM-180, N-CAM-140 and N-CAM-120, where "180", "140" and "120" indicate the apparent molecular weights of the isoforms as measured by SDS polyacrylamide gel electrophoresis) are expressed at least transiently in developing tissues, and permanently in nerve tissue. Both the N-CAM-180 and N-CAM-140 isoforms are expressed in both developing and adult tissue. The N-CAM-120 isoform is found only in adult tissue. Another neural CAM is L1.

The ability of OP-1 based chimeras to stimulate CAM expression can be demonstrated using the following protocol, using NG108-15 cells. NG108-15 is a transformed hybrid cell line (neuroblastoma x glioma, America Type Culture Collection (ATCC), Rockville, MD), exhibiting a morphology characteristic of transformed embryonic neurons. As described in Example D, below, untreated NG108-15 cells exhibit a fibroblastic, or minimally differentiated, morphology and express only the 180 and 140 isoforms of N-CAM normally associated with a developing cell. Following treatment with members of the vg/dpp subgroup these cells exhibit a morphology characteristic of adult neurons and express enhanced levels of all three N-CAM isoforms.

In this example, NG108-15 cells are cultured for 4 days in the presence of increasing concentrations of either the OP-1 morphon or natural OP-1 using standard culturing procedures, and standard Western blots performed on whole cell extracts. N-CAM isoforms are detected with an antibody which crossreacts with all three isoforms, mAb H28.123, obtained from Sigma Chemical Co., St. Louis, the different isoforms being distinguishable by their different mobilities on an electrophoresis gel. Control NG108-15 cells (untreated) express both the 140 kDa and the 180 kDa isoforms, but not the 120 kDa, as determined by Western blot analyses using up to 100 mg of protein. Treatment of NG108-15 cells with OP-1 chimera, like natural OP-1 results in a dose-dependent increase in the expression of the 180 kDa and 140 kDa isoforms, as well as the induction of the 120 kDa isoform induced. In addition, it is contemplated that the OP-1 chimera, like natural OP-1-induced CAM expression may correlate with cell aggregation, as determined by histology.

C. Redifferentiation of Transformed Phenotype.

The OP-1 chimera, like natural OP-1, also induces redifferentiation of transformed cells to a morphology characteristic of untransformed cells. The examples provided below detail morphogen-induced redifferentiation of a transformed human cell line of neuronal origin (NG108-15); as well as mouse neuroblastoma cells (N1E-115), and human embryo carcinoma cells, to a morphology characteristic of untransformed cells.

As described above, NG108-15 is a transformed hybrid cell line produced by fusing neuroblastoma x glioma cells (obtained from ATCC, Rockville, MD), and exhibiting a

morphology characteristic of transformed embryonic neurons, e.g., having a fibroblastic morphology. Specifically, the cells have polygonal cell bodies, short, spike-like processes and make few contacts with neighboring cells. Incubation of NG108-15 cells, cultured in a chemically defined, serum-free medium, with 0.1 to 300 ng/ml of OP-1 chimera or natural OP-1 for four hours induces an orderly, dose-dependent change in cell morphology.

For example, NG108-15 cells are subcultured on poly-L-lysine coated 6 well plates. Each well contains 40-50,000 cells in 2.5 ml of chemically defined medium. On the third day, 2.5 ml of OP-1 chimera or natural OP-1 in 60% ethanol containing 0.025% trifluoroacetic is added to each well. The media is changed daily with new aliquots of morphogen. It is contemplated that OP-1 chimera like OP-1 induces a dose-dependent redifferentiation of the transformed cells, including a rounding of the soma, an increase in phase brightness, extension of the short neurite processes, and other significant changes in the cellular ultrastructure. After several days it is contemplated also that treated cells begin to form epithelioid sheets that then become highly packed, multi-layered aggregates, as determined visually by microscopic examination.

Moreover, it is contemplated that the redifferentiation occur without any associated changes in DNA synthesis, cell division, or cell viability, making it unlikely that the morphologic changes are secondary to cell differentiation or a toxic effect of the morphogen. In addition, it is contemplated that the morphon-induced redifferentiation may not inhibit cell division, as determined by ^{3}H-thymidine uptake, unlike other molecules such as butyrate, DMSO, retinoic acid or Forskolin which have been shown to stimulate differentiation of transformed cells, in analogous experiments. The OP-1 chimera like natural OP-1 maintains cell stability and viability after inducing redifferentiation.

The modified morphogen described herein would, therefore, provide useful therapeutic agents for the treatment of neoplasias and neoplastic lesions of the nervous system, particularly in the treatment of neuroblastomas, including retinoblastomas, and gliomas.

D. Maintenance of Phenotype.

OP-1 chimeras, like natural OP-1 also can be used to maintain a cell's differentiated phenotype. This application is particularly useful for inducing the continued expression of phenotype in senescent or quiescent cells.

1. In Vitro Model for Phenotypic Maintenance.

The phenotypic maintenance capability of morphogens is determined readily. A number of differentiated cells become senescent or quiescent after multiple passages *in vitro* under standard tissue culture conditions well described in the art (e.g., Culture of Animal Cells: A Manual of Basic Techniques, C.R. Freshney, ed., Wiley, 1987). However, if these cells are cultivated *in vitro* in association with a native morphogen such as OP-1, cells are stimulated to maintain expression of their phenotype through multiple passages. For example, the alkaline phosphatase activity of cultured osteoblasts, such as cultured osteosarcoma cells and calvaria cells, is significantly reduced after multiple passages *in vitro*. However, if the cells are cultivated in the presence of OP-1 alkaline phosphatase activity is maintained over extended periods of time. Similarly, phenotypic expression of myocytes also is maintained in the presence of a morphogen. In the experiment, osteoblasts are cultured as described above. The cells are divided into groups, incubated with varying concentrations of either OP-1 chimera or natural OP-1 (e.g., 0-300 ng/ml) and passaged multiple times (e.g., 3-5 times) using standard methodology. Passaged cells then are tested for alkaline phosphatase activity, as described herein as an indication of differentiated cell metabolic function. It is contemplated that osteoblasts cultured in the absence of OP-1 chimera, like natural OP-1, have reduced alkaline phosphatase activity, as compared to OP-1 chimera, or natural OP-1-treated cells.

2. In Vivo Model for Phenotypic Maintenance.

Phenotypic maintenance capability also may be demonstrated *in vivo*, using a standard rat model for osteoporosis. Long Evans female rats (Charles River Laboratories, Wilmington, MA) are Sham-operated (control animals) or ovariectomized using standard surgical techniques, to produce an osteoporotic condition resulting from decreased estrogen production. Following surgery, e.g., 200 days after ovariectomy, rats are systemically provided with phosphate buffered saline (PBS) or morphogen, (e.g., OP-1 chimera, or natural OP-1, 1-10 mg) for 21 days (e.g., by daily tail vein injection.) The rats then are sacrificed and serum alkaline phosphatase levels, serum calcium levels, and serum osteocalcin levels are determined, using standard

methodologies as described therein and above. It is contemplated that the OP-1 chimera treated rats, like the OP-1 treated rats may exhibit elevated levels of osteocalcin and alkaline phosphatase activity. Histomorphometric analysis on the tibial diasypheal bone shows improved bone mass in OP-1 chimera treated animals as compared with untreated, ovariectomized rats.

E. Proliferation of Progenitor Cell Populations.

Progenitor cells can be stimulated to proliferate *in vivo* or *ex vivo*. It is contemplated that cells can be stimulated *in vivo* by injecting or otherwise providing a sterile preparation containing the OP-1 chimera into the individual. For example, the hemopoietic pluripotential stem cell population of an individual can be stimulated to proliferate by injecting or otherwise providing an appropriate concentration of OP-1 chimera to the individual's bone marrow.

Progenitor cells can be stimulated *ex vivo* by contacting progenitor cells of the population to be enhanced with a morphogenically active OP-1 chimera under sterile conditions at a concentration and for a time sufficient to stimulate proliferation of the cells. Suitable concentrations and stimulation times may be determined empirically, essentially following the procedure described in Example A, above. OP-1 chimera concentration of between about 0.1-100 ng/ml and a stimulation period of from about 10 minutes to about 72 hours, or, more generally, about 24 hours, typically should be sufficient to stimulate a cell population of about 10^4 to 10^6 cells. The stimulated cells then may be provided to the individual as, for example, by injecting the cells to an appropriate *in vivo* locus. Suitable biocompatible progenitor cells may be obtained by any of the methods known in the art or described hereinabove.

F. Regeneration of Damaged or Diseased Tissue.

OP-1 chimeras can be used to repair diseased or damaged mammalian tissue. The tissue to be repaired preferably is assessed first, and excess necrotic or interfering scar tissue removed as needed, e.g., by ablation or by surgical, chemical, or other methods known in the medical arts.

OP-1 chimera then can be provided directly to the tissue locus as part of a sterile, biocompatible composition, either by surgical implantation or injection. The chimera also can be provided systemically, as by oral or parenteral administration. Alternatively, a sterile, biocompatible composition containing progenitor cells stimulated by a morphogenically active

OP-1 chimera can be provided to the tissue locus. The existing tissue at the locus, whether diseased or damaged, provides the appropriate matrix to allow the proliferation and tissue-specific differentiation of progenitor cells. In addition, a damaged or diseased tissue locus, particularly one that has been further assaulted by surgical means, provides a morphogenically permissive environment. Systemic provision of OP-1 chimera may be sufficient for certain applications (e.g., in the treatment of osteoporosis and other disorders of the bone remodeling cycle, as an example).

In some circumstances, particularly where tissue damage is extensive, the tissue may not be capable of providing a sufficient matrix for cell influx and proliferation. In these instances, it may be necessary to provide progenitor cells stimulated by the OP-1 chimera to the tissue locus in association with a suitable, biocompatible, formulated matrix, prepared by any of the means described below. The matrix preferably is *in vivo* biodegradable. The matrix also may be tissue-specific and/or may comprise porous particles having dimensions within the range of 70-850mm, most preferably 150-420mm.

OP-1 chimera also may be used to prevent or substantially inhibit immune/inflammatory response-mediated tissue damage and scar tissue formation following an injury. OP-1 chimera may be provided to a newly injured tissue locus, to induce tissue morphogenesis at the locus, preventing the aggregation of migrating fibroblasts into non-differentiated connective tissue. Preferably the OP-1 chimera may be provided as a sterile pharmaceutical preparation injected into the tissue locus within five hours of the injury. Where an immune/inflammatory response is unavoidably or deliberately induced, as part of, for example, a surgical or other aggressive clinical therapy, OP-1 chimera preferably may be provided prophylactically to the patient, prior to, or concomitant with, the therapy.

G. Morphogenesis Assays.

Below are several examples, describing protocols for demonstrating that the chimeric proteins disclosed herein can be used to induce tissue morphogenesis. By way of example, the following is a description of OP-1 chimera-induced tissue morphogenesis in bone, liver, nerve, dentin, cementum and periodontal tissue.

1. OP-1 Chimera-Induced Bone Morphogenesis.

As described earlier, particularly useful mammalian tissue model system for demonstrating and evaluating the morphogenic activity of a protein is the endochondral bone tissue morphogenesis model known in the art and described, for example, in U.S. Pat. No. 4,968,590 and incorporated herein by reference. The ability to induce endochondral bone formation includes the ability to induce proliferation and differentiation of progenitor cells into chondroblasts and osteoblasts, the ability to induce cartilage matrix formation, cartilage calcification, and bone remodeling, and the ability to induce formation of an appropriate vascular supply and hematopoeitic bone marrow differentiation.

The local environment, in which the morphogenic material is placed, is important for tissue morphogenesis. As used herein, "local environment" is understood to include the tissue structural matrix and the environment surrounding the tissue. For example, in addition to needing an appropriate anchoring substratum for their proliferation, the cells stimulated by morphogens need signals to direct the tissue-specificity of their differentiation. These signals vary for the different tissues and may include cell surface markers. In addition, vascularization of new tissue requires a local environment which supports vascularization.

The following sets forth various procedures for evaluating the *in vivo* morphogenic utility of OP-1 chimera and OP-1 chimera containing compositions. The compositions may be injected or surgically implanted in a mammal, following any of a number of procedures well known in the art. For example, surgical implant bioassays may be performed essentially following the procedure of Sampath et al. (1983) Proc. Natl. Acad. Sci. USA 80: 6591-6595 and U.S. Pat No. 4,968,590.

Histological sectioning and staining is preferred to determine the extent of morphogenesis *in vivo*, particularly in tissue repair procedures. Excised implants are fixed in Bouins Solution, embedded in paraffin, and cut into 6-8 mm sections. Staining with toluidine blue or hemotoxylin/eosin demonstrates clearly the ultimate development of the new tissue. Twelve day implants are usually sufficient to determine whether the implants contain newly induced tissue.

Successful implants exhibit a controlled progression through the stages of induced tissue development allowing one to identify and follow the tissue-specific events that occur. For example, in endochondral bone formation the stages include: (1) leukocytes on day one; (2) mesenchymal cell migration and proliferation on days two and three; (3) chondrocyte

appearance on days five and six; (4) cartilage matrix formation on day seven; (5) cartilage calcification on day eight; (6) vascular invasion, appearance of osteoblasts, and formation of new bone on days nine and ten; (7) appearance of osteoclastic cells, and the commencement of bone remodeling and dissolution of the implanted matrix on days twelve to eighteen; and (8) hematopoietic bone marrow differentiation in the resulting ossicles on day twenty-one.

In addition to histological evaluation, biological markers may be used as markers for tissue morphogenesis. Useful markers include tissue-specific enzymes whose activities may be assayed (e.g., spectrophotometrically) after homogenization of the implant. These assays may be useful for quantitation and for rapidly obtaining an estimate of tissue formation after the implants are removed from the animal. For example, alkaline phosphatase activity may be used as a marker for osteogenesis.

Incorporation of systemically provided OP-1 chimera can be followed using labeled protein (e.g., radioactively labelled) and determining its localization in the new tissue, and/or by monitoring their disappearance from the circulatory system using a standard labeling protocol and pulse-chase procedure. OP-1 chimera also may be provided with a tissue-specific molecular tag, whose uptake may be monitored and correlated with the concentration of OP-1 chimera provided. As an example, ovary removal in female rats results in reduced bone alkaline phosphatase activity, and renders the rats predisposed to osteoporosis. If the female rats now are provided with OP-1 chimera, a reduction in the systemic concentration of calcium may be seen, which correlates with the presence of the provided OP-1 chimera and which is anticipated to correspond with increased alkaline phosphatase activity.

2. OP-1 Chimera-Induced Liver Regeneration.

As another example, a method for inducing morphogenesis of substantially injured liver tissue following a partial hepatectomy utilizing OP-1 chimera is presented. Variations on this general protocol may be used to test morphogen activity of OP-1 chimera in other different tissues. The general method involves excising an essentially nonregenerating portion of a tissue and providing OP-1 chimera, preferably as a soluble pharmaceutical preparation to the excised tissue locus, closing the wound, and examining the site at a future date. Like bone, liver has a potential to regenerate upon injury during post-fetal life.

OP-1 chimera, e.g., 1 mg/ml, in a biocompatible solution, for example, a purified OP-1 chimera, is solubilized in 50% ethanol, or compatible solvent, containing 0.1% trifluoroacetic acid, or compatible acid. The injectable OP-1 chimera solution is prepared, e.g., by diluting one volume of OP-1 morphon solvent-acid stock solution with 9 volumes of 0.2% rat serum albumin in sterile PBS (phosphate-buffered saline).

In the experiment, growing rats or aged rats (e.g., Long Evans, Charles River Laboratories, Wilmington, MA) are anesthetized by using ketamine. Two of the liver lobes (left and right) are cut out (approximately 1/3 of the lobe) and the OP-1 chimera is injected locally at multiple sites along the cut ends. The amount of OP-1 chimera injected may be, e.g., 20-100 μg in 100-1000 μl PBS/RSA (phosphate buffered saline/rat serum albumin) injection buffer. Placebo samples comprise injection buffer only. In experimental essays, five rats in each group preferably are used. The wound is closed and the rats are allowed to eat normal food and drink tap water.

After 12 days, the rats are sacrificed and liver regeneration is observed visually, to evaluate the effects of the OP-1 chimera on liver regeneration most effectively. It is contemplated that the OP-1 chimera injected group will show complete liver tissue regeneration with no sign remaining of any cut in the liver. By contrast, the control group into which only PBS is injected, show typically only minimal regeneration with the incision remaining in the sample.

3. OP-1 Chimera-Induced Dentin, Cementum and Periodontal Ligament Regeneration.

As still another example, the ability of OP-1 chimeras to induce dentinogenesis also can be demonstrated. To date, the unpredictable response of dental pulp tissue to injury is a basic clinical problem in dentistry. Cynomolgus monkeys are chosen as primate models as monkeys are presumed to be more indicative of human dental biology than models based on lower non-primate mammals.

Using standard dental surgical procedures, small areas (e.g., 2mm) of dental pulps are surgically exposed by removing the enamel and dentin immediately above the pulp (by drilling) of sample teeth, performing a partial amputation of the coronal pulp tissue, inducing

homeostasis, application of the pulp treatment, and sealing and filling the cavity by standard procedures.

Pulp treatments used include: an OP-1 chimera dispersed in a carrier matrix; carrier matrix alone, and no treatment. Twelve teeth per animal (four for each treatment) are prepared, and two animals are used. At four weeks, teeth are extracted and processed histologically for analysis of dentin formation, and/or ground to analyze dentin mineralization. The effect of OP-1 chimera on osteodentin reparation may be observed visually by comparing with the control samples. It is contemplated that the OP-1 chimera, like natural OP-1, plus a carrier matrix induce formation of reparative, osteodentin bridges, traversing on surgically exposed healthy dental pulps. By contrast, pulps treated with carrier matrix alone, do not form reparative dentin.

4. OP-1 Chimera-Induced Nerve Tissue Repair.

As yet another example, the induction of regenerative effects on central nervous system (CNS) repair, by a morphogenically active OP-1 chimera, can be demonstrated using a rat brain stab model. Briefly, male Long Evans rats are anesthetized and the head area prepared for surgery. The calvariae is exposed using standard surgical procedures and a hole drilled toward the center of each lobe using a 0.035K wire, just piercing the calvariae. 25ml solutions containing either OP-1 chimera (25mg), natural OP-1 (25mg) or PBS then is provided to each of the holes by Hamilton syringe. Solutions are delivered to a depth approximately 3 mm below the surface, into the underlying cortex, corpus callosum and hippocampus. The skin then is sutured and the animal allowed to recover.

Three days post surgery, rats are sacrificed by decapitation and their brains processed for sectioning. Scar tissue formation is evaluated by immunofluoresence staining for glial fibrillary acidic protein, a marker protein for glial scarring, to qualitatively determine the degree of scar formation. It is contemplated that the OP-1 chimera, like natural OP-1, may result in reduced levels of glial fibrillary acidic protein in the tissue sections of animals treated with OP-1 chimera, evidencing the ability of the morphogen to inhibit glial scar formation, thereby stimulating nerve regeneration.

The ability of OP-1 chimera to stimulate peripheral nervous system axonal growth over extended distances can be demonstrated using the following model. Neurons of the peripheral nervous system can sprout new processes on their own following injury, but without guidance

these sproutings typically fail to connect appropriately and die. Where the break is extensive, e.g., greater than 5 or 10 mm, regeneration is poor or nonexistent. Previous experiments with OP-1, show that morphogens stimulate peripheral nervous system axonal growth over extended distances, allowing repair and regeneration of damaged peripheral neural pathways.

In this example OP-1 chimera stimulation of nerve regeneration is demonstrated using the rat sciatic nerve model. The rat sciatic nerve can regenerate spontaneously across a 5 mm gap, and occasionally across a 10 mm gap, provided that the severed ends are inserted in a saline-filled nerve guidance channel. In this experiment, nerve regeneration across at least a 12mm gap is tested.

Adult female Sprague-Dawley rats (Charles River Laboratories, Inc.) weighing 230-250 g are anesthetized with intraperitoneal injections of sodium pentobarbital (35 mg/kg body weight). A skin incision is made parallel and just posterior to the femur. The avascular intermuscular plane between vastus lateralis and hamstring muscles are entered and followed to the loose fibroareolar tissue surrounding the sciatic nerve. The loose tissue is divided longitudinally thereby freeing the sciatic nerve over its full extent without devascularizing any portion. Under a surgical microscope the sciatic nerves are transected with microscissors at mid-thigh and grafted with an OP-1 morphon gel graft that separates the nerve stumps by 12 mm. The graft region is encased in a silicone tube 20 mm in length with a 1.5 mm inner diameter, the interior of which is filled with the chimera solution. Specifically, the central 12 mm of the tube consists of an OP-1 morphon gel prepared by mixing 1 to 5 mg of substantially pure OP-1 chimera produced with approximately 100 ml of MATRIGEL^{TM} (from Collaborative Research, Inc., Bedford, MA), an extracellular matrix extract derived from mouse sarcoma tissue, and containing solubilized tissue basement membrane, including laminin, type IV collagen, heparin sulfate, proteoglycan and entactin, in phosphate-buffered saline. The chimera-filled tube then is implanted directly into the defect site, allowing 4 mm on each end to insert the nerve stumps. Each stump is abutted against the chimera gel and is secured in the silicone tube by three stitches of commercially available surgical 10-0 nylon through the epineurium, the fascicle protective sheath.

In addition to OP-1 chimera gel grafts, control grafts of empty silicone tubes, silicone tubes filled with gel only and "reverse" autografts, wherein 12 mm transected segments of the animal's sciatic nerve are rotated 180° prior to suturing, preferably also are grafted. All

experiments preferably are performed in quadruplicate. All wounds preferably are closed by wound clips that are removed after 10 days. Rats can be grafted on both legs. At 3 weeks the animals are sacrificed, and the grafted segments removed and frozen on dry ice immediately. Frozen sections then are cut throughout the graft site, and examined for axonal regeneration by immunofluorescent staining using anti-neurofilament antibodies labeled with fluorescein (obtained, for example, from Sigma Chemical Co., St. Louis).

It is contemplated that regeneration of the sciatic nerve can occur across the entire 12 mm distance in all graft sites when the gap is filled with the OP-1 chimera gel. By contrast, empty silicone tubes, gel alone and reverse autografts do not show nerve regeneration.

IV. GENERAL REMARKS

It will be understood and appreciated that any of the foregoing means for assessing attributes such as folding, solubility, stability, surface binding/adherence, bio-activity, etc., can be performed using chimeric constructs other than OP-1 based constructs. All that is required is a comparative assessment of the native TGF-β superfamily protein and the chimeric construct's attributes. Thus, for example, a chimera which retains the preferred folding attributes of native protein X but which has acquired the neural tissue differentiation attributes of native protein Y can be readily identified using the experimental paradigms and guidance provided herein. The skilled artisan will recognize a preferred construct as one having all the desired attributes optimally combined, with optimal not intended to necessarily mean maximal expression of any one particular attribute. Rather, optimal is intended to mean expression of each desired attribute such that the combined expression is suitable for the circumstances and/or uses contemplated by the artisan.

The invention may be embodied in other specific forms without departing from the spirit and scope thereof. Accordingly, other embodiments are within the following claims.

What is claimed is:

1. A TGF-β superfamily chimeric protein derived from at least two different members of said superfamily, said chimeric protein comprising a dimer wherein one monomer comprises a finger 1 subdomain, a finger 2 subdomain and a heel subdomain, said finger 2 subdomain being derived from a first member of said superfamily, said finger 1 or heel subdomain being derived from a second, different member of said superfamily, wherein said monomer further comprises a conserved C-terminal cysteine skeleton.

2. The chimeric protein of claim 1, wherein said second member is OP-1 (SEQ ID NO: 39).

3. The chimeric protein of claim 1, wherein said first member is CDMP-2 (SEQ ID NO: 10) and said second member is OP-1 (SEQ ID NO: 39).

4. The chimeric protein of claim 1, wherein said first member is BMP-2 (SEQ ID NO: 2) and said second member is OP-1 (SEQ ID NO: 39).

5. The chimeric protein of claim 1, wherein said finger 1 subdomain is derived from OP-1 (SEQ ID NO: 39), said finger 2 subdomain is derived from CDMP-2 (SEQ ID NO: 10), and said heel domain comprises at least a portion of the heel domain of OP-1 (SEQ ID NO: 39).

6. The chimeric protein of claim 1, wherein said protein comprises a dimer having identical monomers.

7. The chimeric protein of claim 1, wherein said monomer further comprises a finger 1 or heel subdomain derived from a third, different member of said superfamily.

8. A DNA sequence encoding the monomer of said TGF-β superfamily chimeric protein of claim 1.

9. A method of manufacturing a TGF-β superfamily chimeric protein having an altered attribute relative to a native TGF-β superfamily protein, said method comprising the steps of:

(a) producing the chimeric protein of claim 1.

10. The method of claim 9, wherein said altered attribute is selected from the group consisting of: altered amino acid sequence; altered protein folding; altered protein stability; altered protein solubility; altered isoelectric point; altered binding to solid surfaces; altered binding to solubilized carriers; altered biospecificity; altered specific activity; and altered morphogenic activity.

11. The method of claim 9, wherein said altered attribute is altered amino acid sequence.

12. The method of claim 9, wherein said chimeric protein has a plurality of altered attributes.

13. The method of claim 9, wherein step (a) is accomplished using a production method selected from the group consisting of: automated peptide synthesis (chemosynthesis); expressing a DNA sequence encoding said chimeric protein (recombinant synthesis); and ligating peptide fragments corresponding to one or more subdomains or fragments thereof (biosynthesis).

14. The method of claim 9, wherein step (a) is accomplished by providing conditions under which protein folding can occur.

15. A method of inducing tissue morphogenesis, said method comprising the step of: administering a morphogenic effective amount of the chimeric protein of claim 1, 2, 3, 4, 5 or 6.

16. The method of claim 15, wherein said tissue is selected from the group consisting of: bone; non-mineralized skeletal tissue; dental tissue; connective tissue; brain; liver; and nerve.

17. A method of determining the epitope recognized by an antibody that binds a first TGF-β superfamily protein and does not bind a second TGF-β superfamily protein, the method comprising:

replacing part or all of a C-terminal subdomain of the first TGF-β superfamily protein with part or all of the corresponding subdomain of the second TGF-β superfamily protein to create a chimeric protein;

exposing the chimeric protein to the antibody under conditions that allow folding of the chimeric protein and binding of the antibody to its antigen; and

detecting the binding of the antibody to the chimeric protein, wherein presence of the binding indicates that the epitope of the antibody does not reside within the replaced portion of the C-terminal subdomain of the first TGF-β superfamily protein, and absence of the binding indicates that the epitope resides within the replaced portion of the C-terminal domain of the first TGF-β superfamily protein.

Abstract of the Disclosure

The invention provides modified proteins and DNAs of the TGF-β superfamily including modified morphogenic proteins. The proteins of the present invention display altered biological or biochemical attributes. Specifically, the modified proteins are designed through manipulation or exchange of subdomains among two or more members of the superfamily such that certain desired attributes are mixed-and-matched and then exploited therapeutically.

542MKK7557/47.A770619v2

Fig. 1

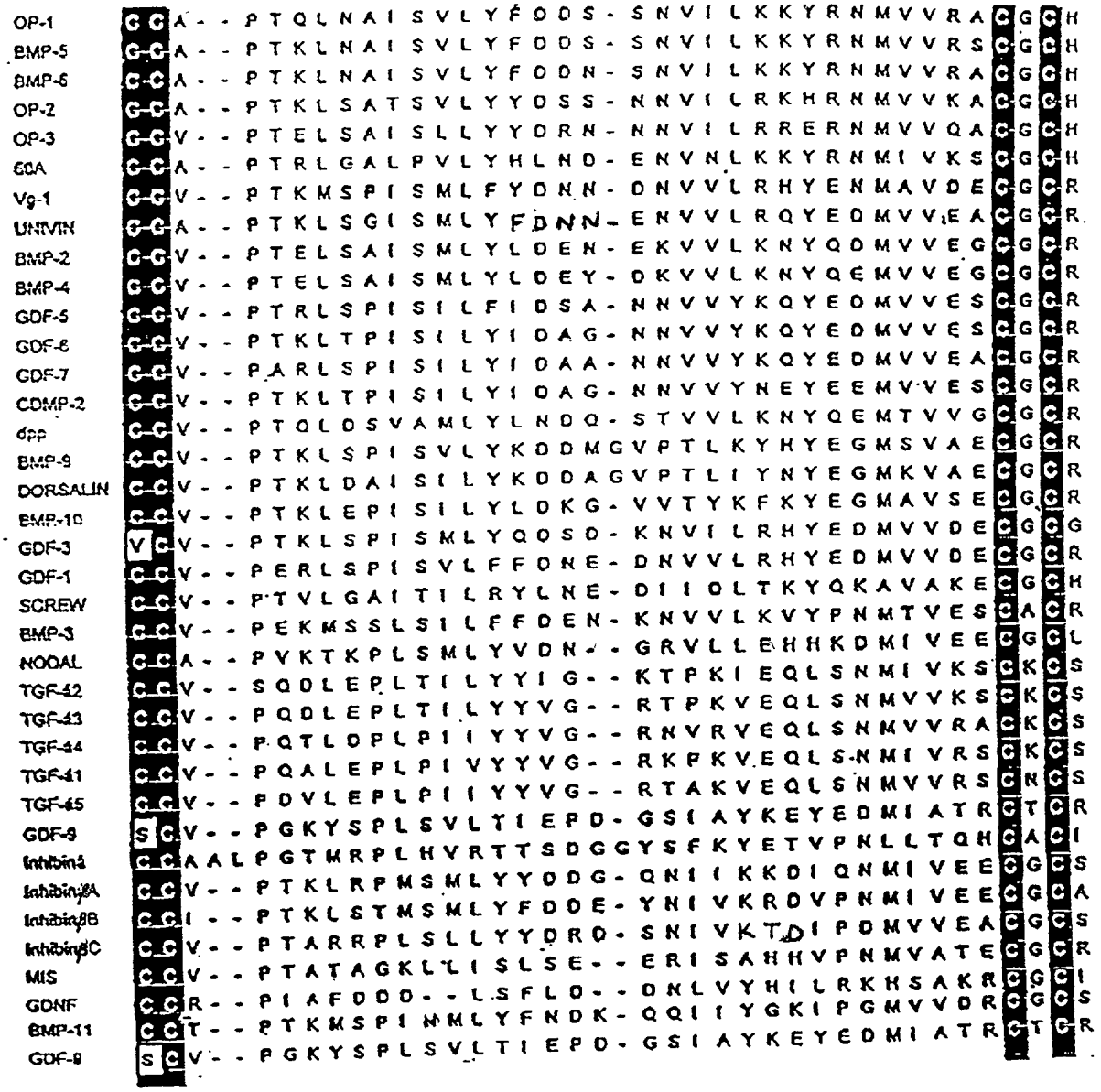
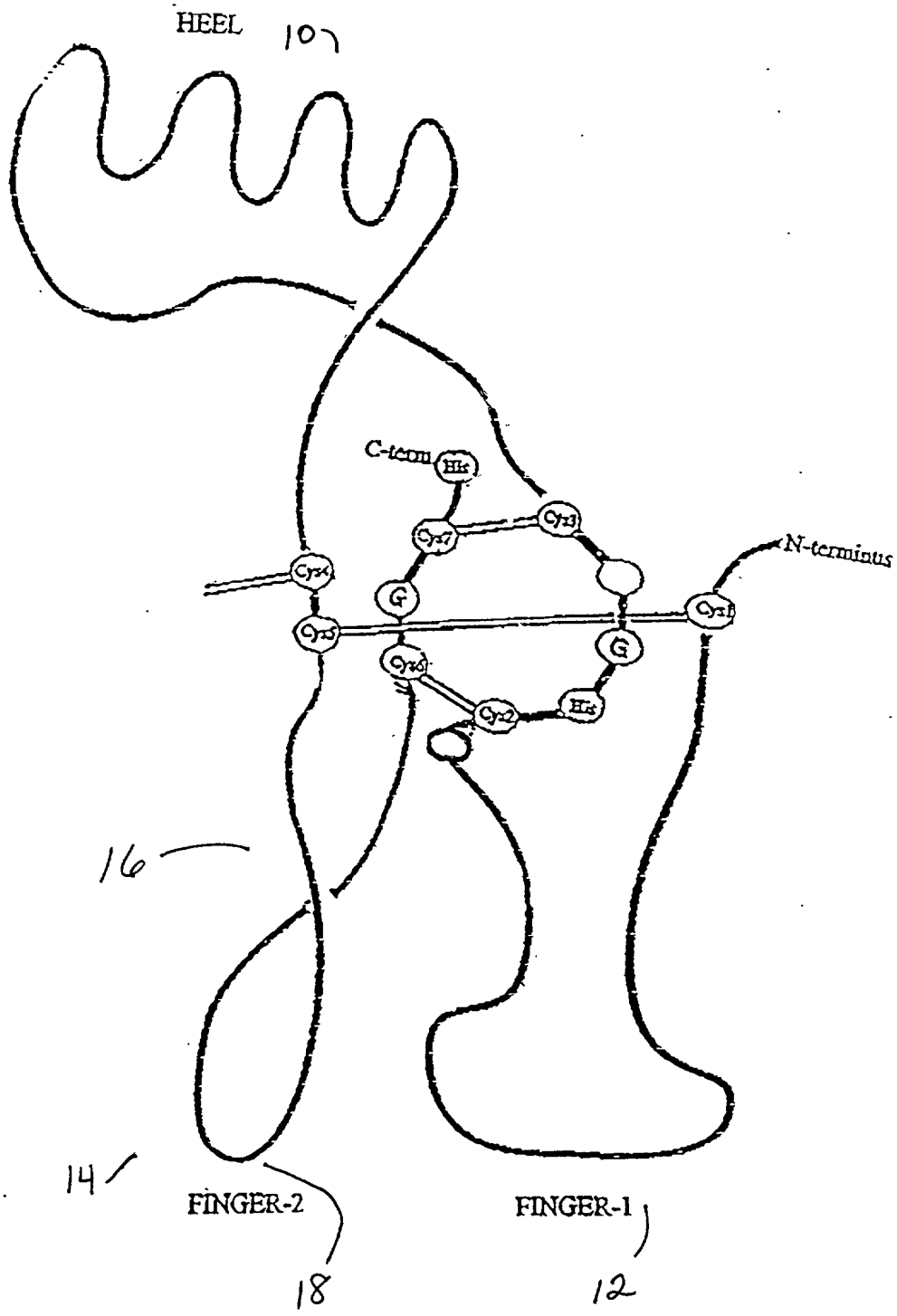

OP-1
BMP-5
BMP-6
OP-2
OP-3
60A
Vg-1
UNIVIN
BMP-2
BMP-4
GDF-5
GDF-6
GDF-7
CDMP-2
dpp
BMP-9
DORSALIN
BMP-10
GDF-3
GDF-1
SCREW
BMP-3
NODAL
TGF-β2
TGF-β3
TGF-β4
TGF-β1
TGF-β5
GDF-9
Inhibin α
Inhibin βA
Inhibin βB
Inhibin βC
MIS
GDNF
BMP-11
GDF-8

Fig. 2


HEEL
10
C-term
His
Cys7
Cys3
Cys4
N-terminus
G
Cys1
Cys5
G
Cys6
Cys2
His
16
14
FINGER-2
FINGER-1
18
12

66918O"9S6h/S6O

7-cysteine domain of OP-1

finger-1

```
TGTAAGAAGCACGAGCTGTATGTCAGCTTCCGAGACCTGGGCTGGCAGGACTGGATCATCGCGCCTGAAGGCTACGCCGCCTACTACTGTGAGGGG
 C  K  K  H  E  L  Y  V  S  F  R  D  L  G  W  Q  D  W  I  I  A  P  E  G  Y  A  A  Y  Y  C  E  G
```

heel

```
GAGTGTGCCTTCCCTCTGAACTCCTACATGAACGCCACCAACCACGCCATCGTGCAGACGCTGGTCCACTTCATCAACCCGGAAACGGTGCCCAAGCCCTGC
 E  C  A  F  P  L  N  S  Y  M  N  A  T  N  H  A  I  V  Q  T  L  V  H  F  I  N  P  E  T  V  P  K  P  C
```

finger-2

```
TGTGCGCCCACGCAGCTCAATGCCATCTCCGTCCTCTACTTCGATGACAGCTCCAACGTCATCCTGAAGAAATACAGAAACATGGTGGTCCGGGCCTGTGGCTGCCAC
 C  A  P  T  Q  L  N  A  I  S  V  L  Y  F  D  D  S  S  N  V  I  L  K  K  Y  R  N  M  V  V  R  A  C  G  C  H
```

Fig. 3

Figure 4 A
OP-1 chimerics with CDMP-2 or with BMP-2

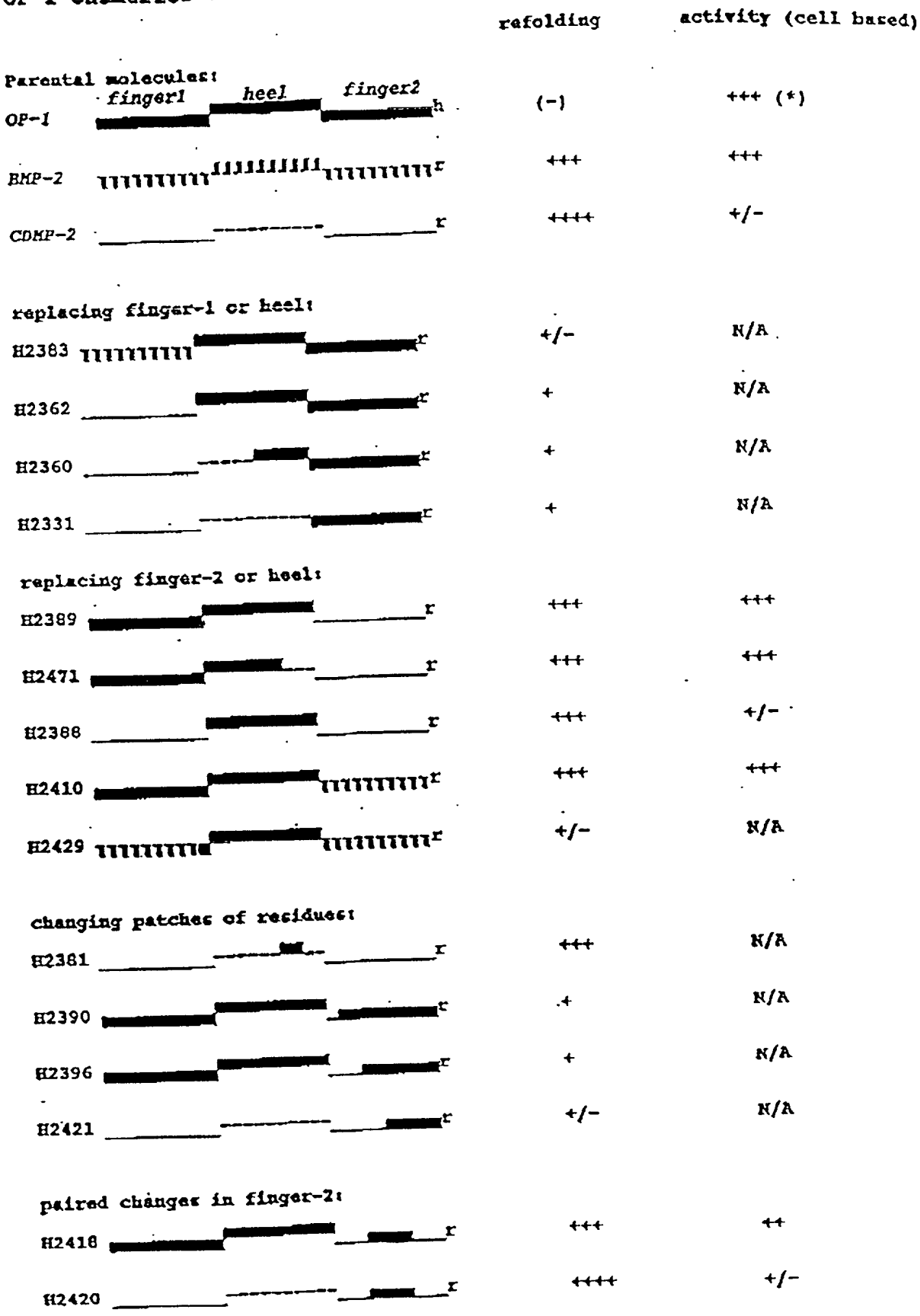
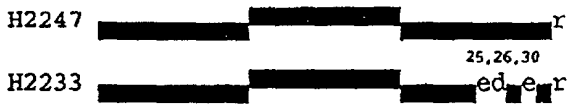
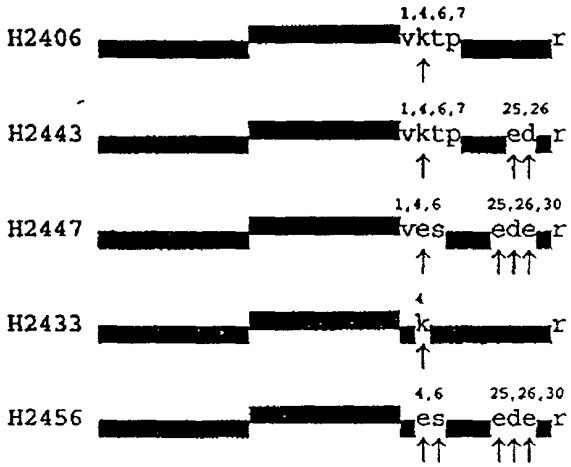


Figure 4B

OP-1 mutants with C-terminal arginine instead of histidine:



Balancing of charged residues in finger-2 of OP-1 mutants:



Correlation of Refolding Efficiency and Charged Amino Acids
in the TGF-β (Seven Cysteine) Domain

protein	finger-1	CXGXC	heel	finger-2	CXCX C-term	Total of charged residues (+), (-) = total	negative charges, finger-2	net charges, finger-2	refolding efficiency
OP-1	3+, 4-	2-	1+, 1-	4+, 2-	0	8+, 9- = 17	2-	2+	+/-
H2247	3+, 4-	2-	1+, 1-	4+, 2-	1+	9+, 9- = 18	2-	2+	+
H2447	3+, 4-	2-	1+, 1-	2+, 6-	1+	7+, 12- = 19	6-	4-	+++
BMP-3	4+, 4-	0	3+, 1-	3+, 4-	1+	11+, 9- = 20	4-	1-	+++
BMP-2	2+, 3-	1-	2+, 1-	2+, 6-	1+	7+, 11- = 18	6-	4-	+++
GDF-5	3+, 5-	1-	1+, 4-	2+, 4-	1+	6+, 14- = 20	4-	2-	+++
CDMP-2	3+, 5-	1-	1+, 3-	2+, 4-	1+	6+, 13- = 19	4-	2-	+++
GDNF	2+, 4-	0	6+, 4-	5+, 5-	0	13+, 13- = 26	5-	0	+++
TGF-β1	5+, 3-	0	1+, 1-	5+, 2-	1+	11+, 6- = 17	2-	3+	+/-
TGF-β2	5+, 3-	0	1+, 2-	4+, 3-	1+	10+, 8- = 18	3-	1+	+/-

Figure 5

3536-4536-5

66918D"9C6h/C6D

FIG. 6A

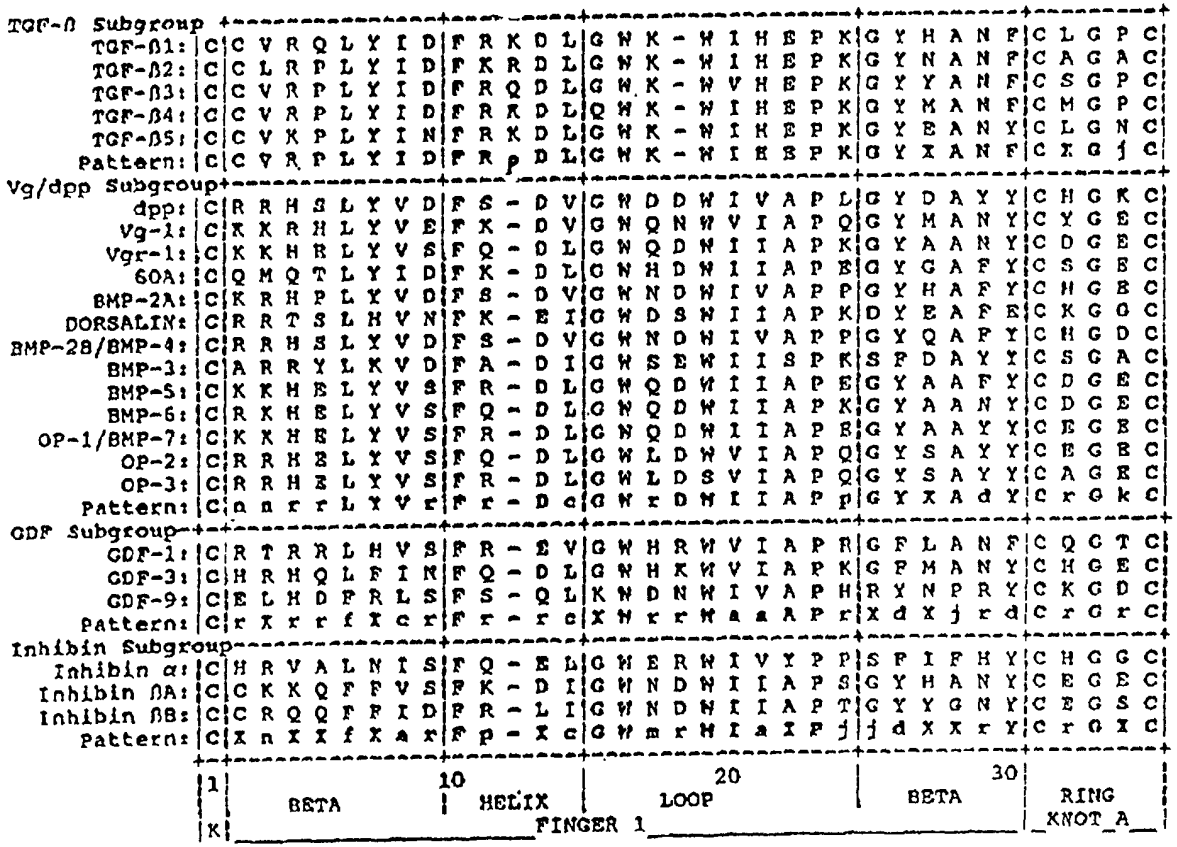


FIG. 6B

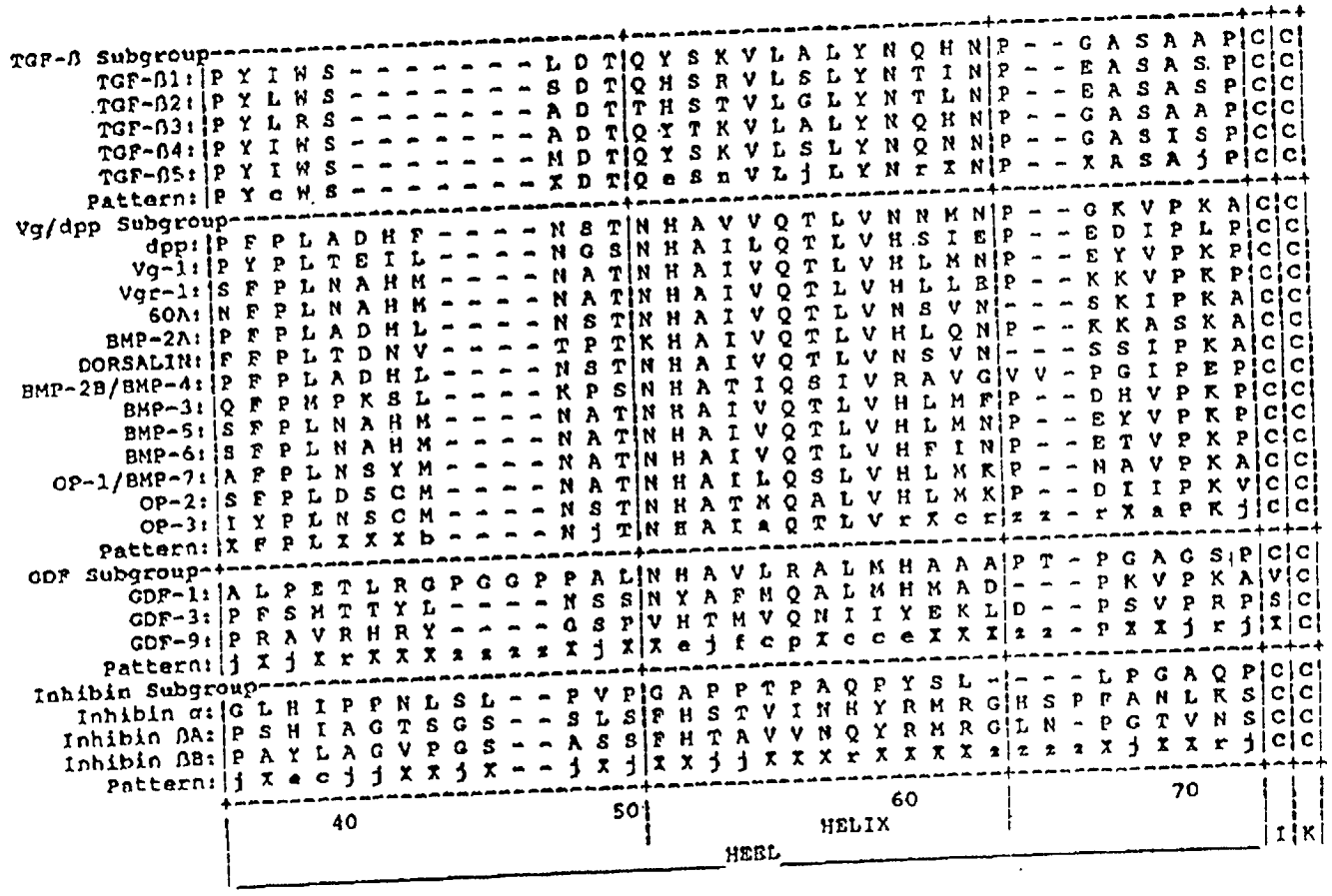

TGF-ß Subgroup
TGF-ß1:
TGF-ß2:
TGF-ß3:
TGF-ß4:
TGF-ß5:
Pattern:
Vg/dpp Subgroup
dpp:
Vg-1:
Vgr-1:
60A:
BMP-2A:
DORSALIN:
BMP-2B/BMP-4:
BMP-3:
BMP-5:
BMP-6:
OP-1/BMP-7:
OP-2:
OP-3:
Pattern:
GDF Subgroup
GDF-1:
GDF-3:
GDF-9:
Pattern:
Inhibin Subgroup
Inhibin α:
Inhibin ßA:
Inhibin ßB:
Pattern:
40
50
60
70
HELIX
HEEL
I
K

FIG. 6C

```
TGF-ß subgroup-------------------------+------------------+-------------------------+-------+-+
     TGF-ß1:|V - - P Q A L E P L P I V Y|Y V G - - R K P|K V E Q L S N M I V R S|C K C|S|
     TGF-ß2:|V - - S Q D L E P L T I L Y|Y I G - - K T P|K I E Q L S N M I V K S|C K C|S|
     TGF-ß3:|V - - P Q D L E P L T I L Y|Y V G - - R T P|K V E Q L S N M V V K S|C K C|S|
     TGF-ß4:|V - - P Q T L D P L P I I Y|Y V G - - R N V|R V E Q L S N M V V R A|C K C|S|
     TGF-ß5:|V - - P D V L E P L P I I Y|Y V G - - R T A|K V E Q L S N M V V R S|C N C|S|
    Pattern:|V - - P Q X L E P L j I c Y|Y V G - - R r j|K V E Q L S N M a V o S|C K C|S|
Vg/dpp Subgroup------------------------+------------------+-------------------------+-------+-+
        dpp:|V - - P T Q L D S V A M L Y|L N D Q - S T V|V L K N Y Q E M T V V G|C G C|R|
       Vg-1:|V - - P T K M S P I S M L F|Y D N N - D N V|V L R H Y E N M A V D E|C G C|R|
      Vgr-1:|A - - P T K L N A I S V L Y|F D D N - S N V|I L K K Y R N M V V R A|C G C|H|
        60A:|A - - P T R L G A L P V L Y|H L N D - E N V|N L K K Y R N M I V K S|C G C|H|
     BMP-2A:|V - - P T E L S A I S M L Y|L D E N - E K V|V L K N Y Q D M V V E G|C G C|R|
   DORSALIN:|V - - P T K L D A I S I L Y|K D D A G V P T|L I Y N Y E G M K V A E|C G C|R|
BMP-2B/BMP-4:|V - - P T E L S A I S M L Y|L D E Y - D K V|V L K N Y Q E M V V E G|C G C|R|
      BMP-3:|V - - P E K M S S L S I L F|F D E N - K N V|V L K V Y P N M T V E S|C A C|R|
      BMP-5:|A - - P T K L N A I S V L Y|F D D S - S N V|I L K K Y R N M V V R S|C G C|H|
      BMP-6:|A - - P T K L N A I S V L Y|F D D N - S N V|I L K K Y R N M V V R A|C G C|H|
 OP-1/BMP-7:|A - - P T Q L N A I S V L Y|F D D S - S N V|I L K K Y R N M V V R A|C G C|H|
       OP-2:|A - - P T K L S A T S V L Y|Y D S S - N N V|I L R K H R N M V V K A|C G C|H|
       OP-3:|V - - P T E L S A I S L L Y|Y D R N - N N V|I L R R E R N M V V Q A|C G C|H|
    Pattern:|X - - P T p L r A a S c L Y|f D m r z r r V|a L n r Y p l M X V p j|C G C|r|
GDF Subgroup-+-------------------------+------------------+-------------------------+-------+-+
      GDF-1:|V - - P E R L S P I S V L F|F D N S - D N V|V L R H Y E D M V V D E|C G C|R|
      GDF-3:|V - - P T K L S P I S M L Y|Q D S D - K N V|I L R H Y E D M V V D E|C G C|G|
      GDF-9:|V - - P G K Y S P L S V L T|I E P D - G S I|A Y K E Y E D M I A T R|C T C|R|
    Pattern:|V - - P X n f S P o S c L X|X k X r - X r a|X f n r Y E D M a X r p|C j C|X|
Inhibin Subgroup-----------------------+------------------+-------------------------+-------+-+
  Inhibin α:|A A L P G T M R P L H V R T|T S D G G Y S F|K Y E T V P N L L T Q H|C A C|I|
 Inhibin ßA:|V - - P T K L R P M S M L Y|Y D D G - Q N I|I K K D I Q N M I V E E|C G C|S|
 Inhibin ßB:|I - - P T K L S T M S M L Y|F D D E - Y N I|V K R D V P N M I V E E|C G C|A|
    Pattern:|X z z P j r b r j b r o X X|X r D X z X r f|X X p r a X N b c X o r|C h C|X|
            +--------------------------+------------------+-------------------------+-------+-+
            |           80             | 90               |   100                   |110    | |
            |          BETA            |      LOOP        |          BETA           | RING  | |
            |                          FINGER_2_____________________________________|KNOT2  |C|
            |__________________________|
```

669780 9264260

7

FIG.

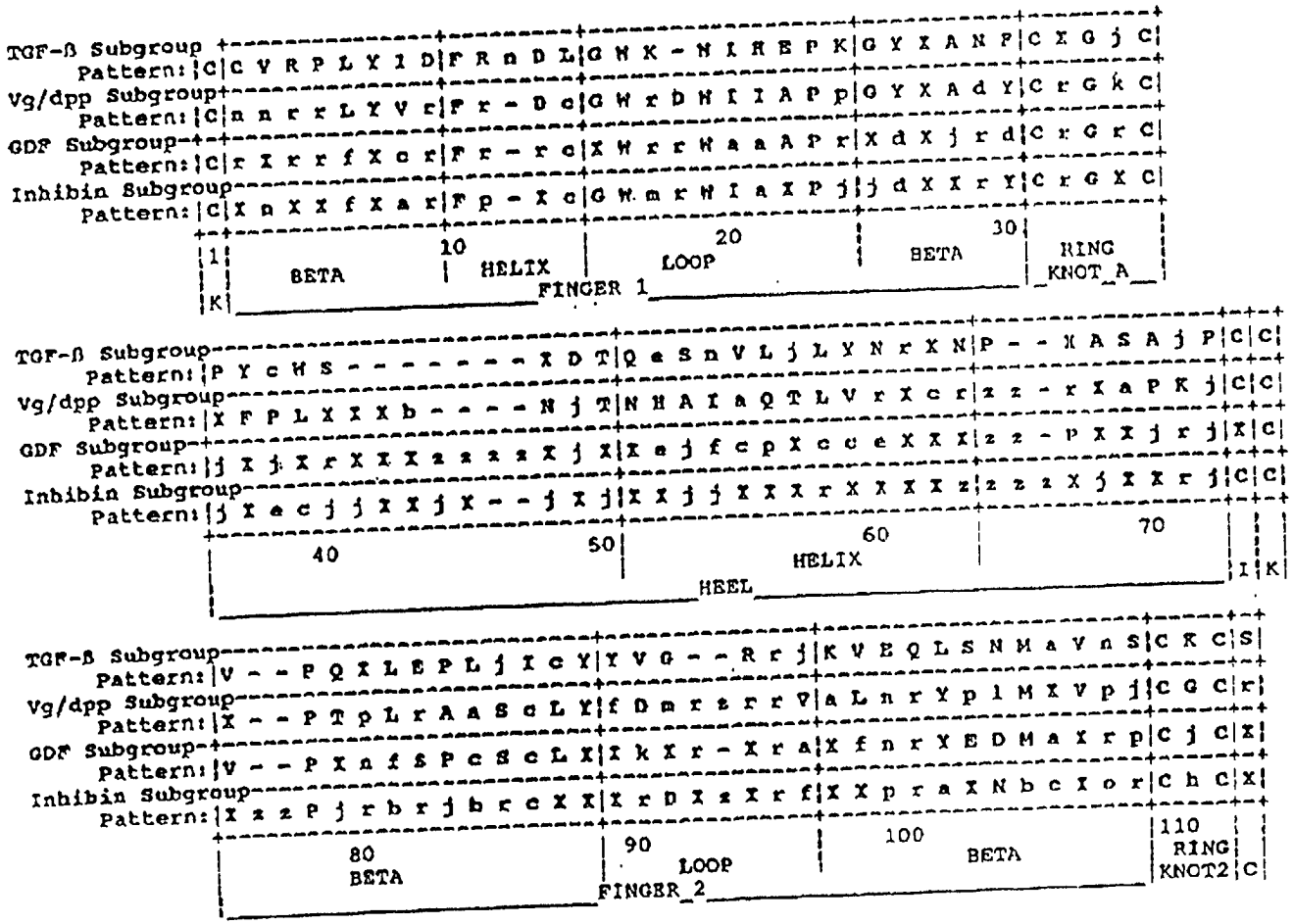
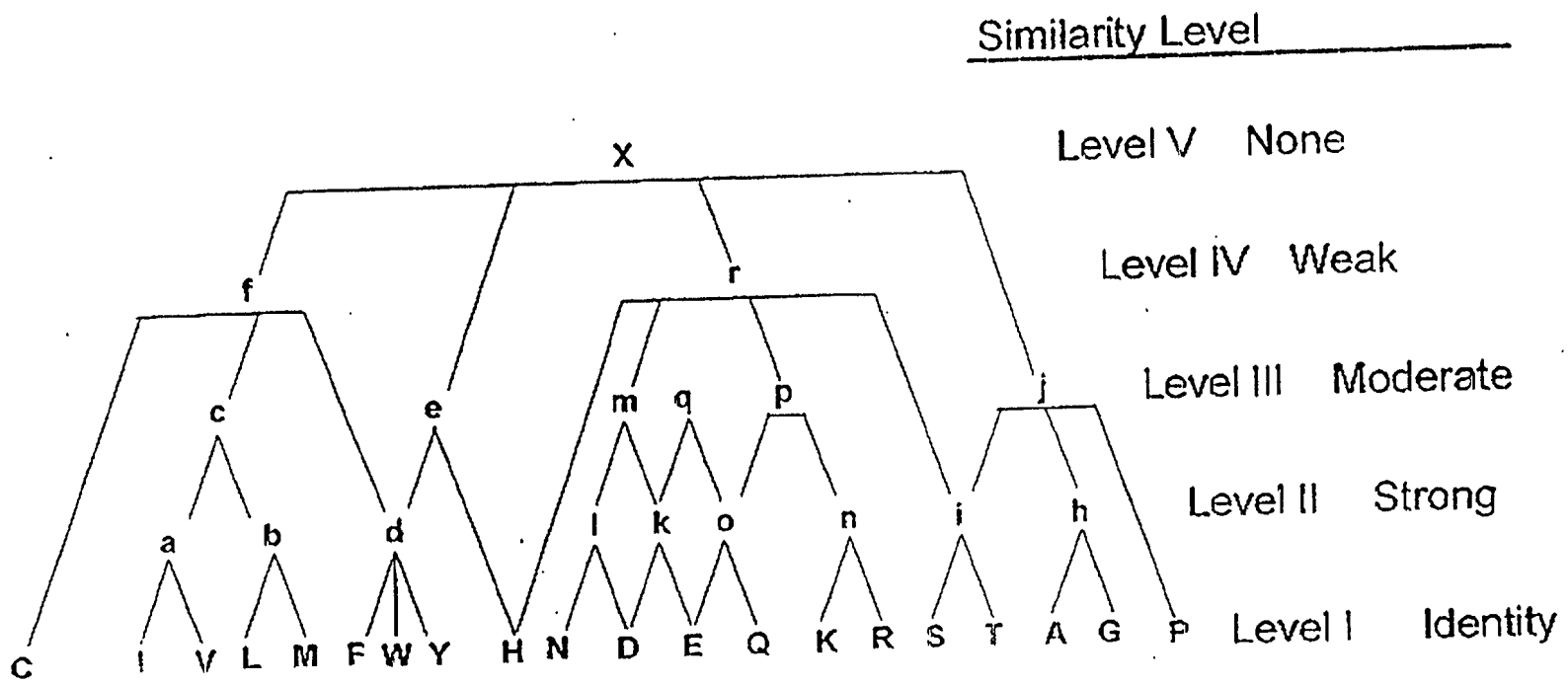


FIG. 8


Similarity Level
Level V None
Level IV Weak
Level III Moderate
Level II Strong
Level I Identity
X
f
r
c
e
m
q
p
j
a
b
d
l
k
o
n
i
h
C
I
V
L
M
F
W
Y
H
N
D
E
Q
K
R
S
T
A
G
P

EXPRESS MAIL LABEL NO. EL280661249US

PATENT
Attorney Docket No. STK-077

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT(S):	Hermann Oppermann et al.		
SERIAL NO.:	Not yet assigned	GROUP NO.:	Not yet assigned
FILING DATE:	June 25, 1999	EXAMINER:	Not yet assigned
TITLE:	MODIFIED PROTEINS AND DNAs OF THE TFG-β SUPERFAMILY, INCLUDING MODIFIED MORPHOGENIC PROTEINS		

BOX SEQUENCE
Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

SUBMISSION OF PAPER COPY AND/OR COMPUTER READABLE COPY OF SEQUENCE LISTING FOR INVENTION CONTAINING NUCLEOTIDE AND/OR AMINO ACID SEQUENCE

(check and complete this item, if applicable)

1. ☐ This replies to the Patent Office Letter dated _____

Note: If these papers are filed before the office letter issues adequate identification of the original papers should be made, e.g., in addition to the name of the inventor and title of invention, the filing date based on the "Express Mail" procedure, the serial number from the return post card or the attorney's docket number added.

☐ A copy of the Patent Office Letter is enclosed.

2. Submitted herewith is/are
(check each item as applicable)

A. ☒ a paper copy of the Sequence Listing for this application with each sequence assigned a separate identifier.

B. ☒ a copy, in computer readable form, of the Sequence Listing for this application.

STATEMENT

3. I hereby state that:

(complete applicable items A, B and/or C)

A. ☒ the content of the paper and computer readable copies submitted herewith are the same.

B. ☐ the content of the computer readable copy submitted herewith is the same as the Sequence Listing appearing on pages __ to __ of the original specification as filed.

C. ☒ this submission includes no new matter.

Respectfully submitted,

Date: August 16, 1999
Reg. No. 40,702

Timothy P. Linkkila
Attorney for Applicants
Testa, Hurwitz, & Thibeault, LLP
High Street Tower
125 High Street
Boston, Massachusetts 02110

Tel. No.: (617) 248-7045
Fax No.: (617) 248-7100

VITO7557/47.A814050_1

SEQUENCE LISTING

<110> Oppermann, Hermann
Tai, Mei-Sheng
McCartney, John

<120> Modified TGF-beta Superfamily Proteins

<130> STK-077

<140>
<141>

<160> 88

<170> PatentIn Ver. 2.0

<210> 1
<211> 35
<212> PRT
<213> Drosophila melanogaster

<220>
<223> 60-A

<400> 1
Ala Pro Thr Arg Leu Gly Ala Leu Pro Val Leu Tyr His Leu Asn Asp
1 5 10 15

Glu Asn Val Asn Leu Lys Lys Tyr Arg Asn Met Ile Val Lys Ser Cys
20 25 30

Gly Cys His
35

<210> 2
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<223> BMP-2

<400> 2
Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp Glu Asn
1 5 10 15

Glu Lys Val Val Leu Lys Asn Tyr Gln Asp Met Val Val Glu Gly Cys
20 25 30

Gly Cys Arg
35

<210> 3
<211> 35

<212> PRT
<213> Homo sapiens

<220>
<223> BMP-3

<400> 3
Val Pro Glu Lys Met Ser Ser Leu Ser Ile Leu Phe Phe Asp Glu Asn
1 5 10 15

Lys Asn Val Val Leu Lys Val Tyr Pro Asn Met Thr Val Glu Ser Cys
20 25 30

Ala Cys Arg
35

<210> 4
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<223> BMP-4

<400> 4
Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp Glu Tyr
1 5 10 15

Asp Lys Val Val Leu Lys Asn Tyr Gln Glu Met Val Val Glu Gly Cys
20 25 30

Gly Cys Arg
35

<210> 5
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<223> BMP-5

<400> 5
Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser
1 5 10 15

Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ser Cys
20 25 30

Gly Cys His
35

<210> 6
<211> 35

<212> PRT
<213> Homo sapiens

<220>
<223> BMP-6

<400> 6
Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Asn
1 5 10 15

Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys
20 25 30

Gly Cys His
35

<210> 7
<211> 36
<212> PRT
<213> Homo sapiens

<220>
<223> BMP-9

<400> 7
Val Pro Thr Lys Leu Ser Pro Ile Ser Val Leu Tyr Lys Asp Asp Met
1 5 10 15

Gly Val Pro Thr Leu Lys Tyr His Tyr Glu Gly Met Ser Val Ala Glu
20 25 30

Cys Gly Cys Arg
35

<210> 8
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<223> BMP-10

<400> 8
Val Pro Thr Lys Leu Glu Pro Ile Ser Ile Leu Tyr Leu Asp Lys Gly
1 5 10 15

Val Val Thr Tyr Lys Phe Lys Tyr Glu Gly Met Ala Val Ser Glu Cys
20 25 30

Gly Cys Arg
35

<210> 9
<211> 35

<212> PRT
<213> Homo sapiens

<220>
<223> BMP-11

<400> 9
Thr Pro Thr Lys Met Ser Pro Ile Asn Met Leu Tyr Phe Asn Asp Lys
1 5 10 15

Gln Gln Ile Ile Tyr Gly Lys Ile Pro Gly Met Val Val Asp Arg Cys
20 25 30

Gly Cys Ser
35

<210> 10
<211> 35
<212> PRT
<213> Bos taurus

<220>
<223> CDMP-2

<400> 10
Val Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile Asp Ala Gly
1 5 10 15

Asn Asn Val Val Tyr Asn Glu Tyr Glu Glu Met Val Val Glu Ser Cys
20 25 30

Gly Cys Arg
35

<210> 11
<211> 36
<212> PRT
<213> Gallus gallus

<220>
<223> Dorsalin

<400> 11
Val Pro Thr Lys Leu Asp Ala Ile Ser Ile Leu Tyr Lys Asp Asp Ala
1 5 10 15

Gly Val Pro Thr Leu Ile Tyr Asn Tyr Glu Gly Met Lys Val Ala Glu
20 25 30

Cys Gly Cys Arg
35

<210> 12
<211> 35

<212> PRT
<213> Drosophila melanogaster

<220>
<223> DPP

<400> 12
Val Pro Thr Gln Leu Asp Ser Val Ala Met Leu Tyr Leu Asn Asp Gln
1 5 10 15

Ser Thr Val Val Leu Lys Asn Tyr Gln Glu Met Thr Val Val Gly Cys
20 25 30

Gly Cys Arg
35

<210> 13
<211> 35
<212> PRT
<213> Mus musculus

<220>
<223> GDF-1

<400> 13
Val Pro Glu Arg Leu Ser Pro Ile Ser Val Leu Phe Phe Asp Asn Glu
1 5 10 15

Asp Asn Val Val Leu Arg His Tyr Glu Asp Met Val Val Asp Glu Cys
20 25 30

Gly Cys Arg
35

<210> 14
<211> 35
<212> PRT
<213> Mus musculus

<220>
<223> GDF-3

<400> 14
Val Pro Thr Lys Leu Ser Pro Ile Ser Met Leu Tyr Gln Asp Ser Asp
1 5 10 15

Lys Asn Val Ile Leu Arg His Tyr Glu Asp Met Val Val Asp Glu Cys
20 25 30

Gly Cys Gly
35

<210> 15
<211> 35

<212> PRT
<213> Homo sapiens

<220>
<223> GDF-5

<400> 15
Val Pro Thr Arg Leu Ser Pro Ile Ser Ile Leu Phe Ile Asp Ser Ala
1 5 10 15

Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ser Cys
20 25 30

Gly Cys Arg
35

<210> 16
<211> 35
<212> PRT
<213> Mus musculus

<220>
<223> GDF-6

<400> 16
Val Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile Asp Ala Gly
1 5 10 15

Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ser Cys
20 25 30

Gly Cys Arg
35

<210> 17
<211> 35
<212> PRT
<213> Mus musculus

<220>
<223> GDF-7

<400> 17
Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile Asp Ala Ala
1 5 10 15

Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ala Cys
20 25 30

Gly Cys Arg
35

<210> 18
<211> 35

<212> PRT
<213> Mus musculus

<220>
<223> GDF-9

<400> 18
Val Pro Gly Lys Tyr Ser Pro Leu Ser Val Leu Thr Ile Glu Pro Asp
1 5 10 15

Gly Ser Ile Ala Tyr Lys Glu Tyr Glu Asp Met Ile Ala Thr Arg Cys
20 25 30

Thr Cys Arg
35

<210> 19
<211> 32
<212> PRT
<213> Homo sapiens

<220>
<223> GDNF

<400> 19
Arg Pro Ile Ala Phe Asp Asp Asp Leu Ser Phe Leu Asp Asp Asn Leu
1 5 10 15

Val Tyr His Ile Leu Arg Lys His Ser Ala Lys Arg Cys Gly Cys Ile
20 25 30

<210> 20
<211> 38
<212> PRT
<213> Homo sapiens

<220>
<223> Inhibin Alpha

<400> 20
Ala Ala Leu Pro Gly Thr Met Arg Pro Leu His Val Arg Thr Thr Ser
1 5 10 15

Asp Gly Gly Tyr Ser Phe Lys Tyr Glu Thr Val Pro Asn Leu Leu Thr
20 25 30

Gln His Cys Ala Cys Ile
35

<210> 21
<211> 35

<212> PRT
<213> Homo sapiens

<220>
<223> Inhibin BetaA

<400> 21
Val Pro Thr Lys Leu Arg Pro Met Ser Met Leu Tyr Tyr Asp Asp Gly
1 5 10 15

Gln Asn Ile Ile Lys Lys Asp Ile Gln Asn Met Ile Val Glu Glu Cys
20 25 30

Gly Cys Ser
35

<210> 22
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<223> Inhibin BetaB

<400> 22
Ile Pro Thr Lys Leu Ser Thr Met Ser Met Leu Tyr Phe Asp Asp Glu
1 5 10 15

Tyr Asn Ile Val Lys Arg Asp Val Pro Asn Met Ile Val Glu Glu Cys
20 25 30

Gly Cys Ala
35

<210> 23
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<223> Inhibin BetaC

<400> 23
Val Pro Thr Ala Arg Arg Pro Leu Ser Leu Leu Tyr Tyr Asp Arg Asp
1 5 10 15

Ser Asn Ile Val Lys Thr Asp Ile Pro Asp Met Val Val Glu Ala Cys
20 25 30

Gly Cys Ser
35

<210> 24
<211> 34

<212> PRT
<213> Homo sapiens

<220>
<223> MIS

<400> 24
Val Pro Thr Ala Tyr Ala Gly Lys Leu Leu Ile Ser Leu Ser Glu Glu
1 5 10 15

Arg Ile Ser Ala His His Val Pro Asn Met Val Ala Thr Glu Cys Gly
20 25 30

Cys Arg

<210> 25
<211> 34
<212> PRT
<213> Mus musculus

<220>
<223> Nodal

<400> 25
Ala Pro Val Lys Thr Lys Pro Leu Ser Met Leu Tyr Val Asp Asn Gly
1 5 10 15

Arg Val Leu Leu Glu His His Lys Asp Met Ile Val Glu Glu Cys Gly
20 25 30

Cys Leu

<210> 26
<211> 35
<212> PRT
<213> Homo sapiens

<220>
<223> OP-2

<400> 26
Ala Pro Thr Lys Leu Ser Ala Thr Ser Val Leu Tyr Tyr Asp Ser Ser
1 5 10 15

Asn Asn Val Ile Leu Arg Lys His Arg Asn Met Val Val Lys Ala Cys
20 25 30

Gly Cys His
35

<210> 27
<211> 35

<212> PRT
<213> Mus musculus

<220>
<223> OP-3

<400> 27
Val Pro Thr Glu Leu Ser Ala Ile Ser Leu Leu Tyr Tyr Asp Arg Asn
1 5 10 15

Asn Asn Val Ile Leu Arg Arg Glu Arg Asn Met Val Val Gln Ala Cys
20 25 30

Gly Cys His
35

<210> 28
<211> 35
<212> PRT
<213> Drosophila melanogaster

<220>
<223> Screw

<400> 28
Val Pro Thr Val Leu Gly Ala Ile Thr Ile Leu Arg Tyr Leu Asn Glu
1 5 10 15

Asp Ile Ile Asp Leu Thr Lys Tyr Gln Lys Ala Val Ala Lys Glu Cys
20 25 30

Gly Cys His
35

<210> 29
<211> 34
<212> PRT
<213> Homo sapiens

<220>
<223> TGF-Beta1

<400> 29
Val Pro Gln Ala Leu Glu Pro Leu Pro Ile Val Tyr Tyr Val Gly Arg
1 5 10 15

Lys Pro Lys Val Glu Gln Leu Ser Asn Met Ile Val Arg Ser Cys Lys
20 25 30

Cys Ser

<210> 30
<211> 34

<212> PRT
<213> Homo sapiens

<220>
<223> TGF-Beta2

<400> 30
Val Ser Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Ile Gly Lys
1 5 10 15

Thr Pro Lys Ile Glu Gln Leu Ser Asn Met Ile Val Lys Ser Cys Lys
20 25 30

Cys Ser

<210> 31
<211> 34
<212> PRT
<213> Homo sapiens

<220>
<223> TGF-Beta3

<400> 31
Val Pro Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Val Gly Arg
1 5 10 15

Thr Pro Lys Val Glu Gln Leu Ser Asn Met Val Val Lys Ser Cys Lys
20 25 30

Cys Ser

<210> 32
<211> 34
<212> PRT
<213> Gallus gallus

<220>
<223> TGF-Beta4

<400> 32
Val Pro Gln Thr Leu Asp Pro Leu Pro Ile Ile Tyr Tyr Val Gly Arg
1 5 10 15

Asn Val Arg Val Glu Gln Leu Ser Asn Met Val Val Arg Ala Cys Lys
20 25 30

Cys Ser

<210> 33
<211> 34

<212> PRT
<213> Xenopus laevis

<220>
<223> TGF-Beta5

<400> 33
Val Pro Asp Val Leu Glu Pro Leu Pro Ile Ile Tyr Tyr Val Gly Arg
1 5 10 15

Thr Ala Lys Val Glu Gln Leu Ser Asn Met Val Val Arg Ser Cys Asn
20 25 30

Cys Ser

<210> 34
<211> 35
<212> PRT
<213> Strongylocentrotus purpuratus

<220>
<223> UNIVIN

<400> 34
Ala Pro Thr Lys Leu Ser Gly Ile Ser Met Leu Tyr Phe Asp Asn Asn
1 5 10 15

Glu Asn Val Val Leu Arg Gln Tyr Glu Asp Met Val Val Glu Ala Cys
20 25 30

Gly Cys Arg
35

<210> 35
<211> 35
<212> PRT
<213> Xenopus laevis

<220>
<223> VG-1

<400> 35
Val Pro Thr Lys Met Ser Pro Ile Ser Met Leu Phe Tyr Asp Asn Asn
1 5 10 15

Asp Asn Val Val Leu Arg His Tyr Glu Asn Met Ala Val Asp Glu Cys
20 25 30

Gly Cys Arg
35

<210> 36
<211> 21

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:synthetic
primer

<220>
<221> CDS
<222> (1)..(21)

<400> 36
atg tcc acg ggg agc aaa cag 21
Met Ser Thr Gly Ser Lys Gln
1 5

<210> 37
<211> 7
<212> PRT
<213> Artificial Sequence

<400> 37
Met Ser Thr Gly Ser Lys Gln
1 5

<210> 38
<211> 1822
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (49)..(1341)
<223> Morphogenic Protein OP1

<400> 38
ggtgcgggcc cggagcccgg agcccgggta gcgcgtagag ccggcgcg atg cac gtg 57
Met His Val
1

cgc tca ctg cga gct gcg gcg ccg cac agc ttc gtg gcg ctc tgg gca 105
Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala Leu Trp Ala
5 10 15

ccc ctg ttc ctg ctg cgc tcc gcc ctg gcc gac ttc agc ctg gac aac 153
Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser Leu Asp Asn
20 25 30 35

gag gtg cac tcg agc ttc atc cac cgg cgc ctc cgc agc cag gag cgg 201
Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser Gln Glu Arg
40 45 50

cgg gag atg cag cgc gag atc ctc tcc att ttg ggc ttg ccc cac cgc 249
Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu Pro His Arg
55 60 65

```
ccg cgc ccg cac ctc cag ggc aag cac aac tcg gca ccc atg ttc atg   297
Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro Met Phe Met
         70                  75                  80

ctg gac ctg tac aac gcc atg gcg gtg gag gag ggc ggc ggg ccc ggc   345
Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Gly Gly Gly Pro Gly
     85                  90                  95

ggc cag ggc ttc tcc tac ccc tac aag gcc gtc ttc agt acc cag ggc   393
Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser Thr Gln Gly
100                 105                 110                 115

ccc cct ctg gcc agc ctg caa gat agc cat ttc ctc acc gac gcc gac   441
Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr Asp Ala Asp
                120                 125                 130

atg gtc atg agc ttc gtc aac ctc gtg gaa cat gac aag gaa ttc ttc   489
Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys Glu Phe Phe
            135                 140                 145

cac cca cgc tac cac cat cga gag ttc cgg ttt gat ctt tcc aag atc   537
His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu Ser Lys Ile
        150                 155                 160

cca gaa ggg gaa gct gtc acg gca gcc gaa ttc cgg atc tac aag gac   585
Pro Glu Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp
    165                 170                 175

tac atc cgg gaa cgc ttc gac aat gag acg ttc cgg atc agc gtt tat   633
Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Arg Ile Ser Val Tyr
180                 185                 190                 195

cag gtg ctc cag gag cac ttg ggc agg gaa tcg gat ctc ttc ctg ctc   681
Gln Val Leu Gln Glu His Leu Gly Arg Glu Ser Asp Leu Phe Leu Leu
                200                 205                 210

gac agc cgt acc ctc tgg gcc tcg gag gag ggc tgg ctg gtg ttt gac   729
Asp Ser Arg Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu Val Phe Asp
            215                 220                 225

atc aca gcc acc agc aac cac tgg gtg gtc aat ccg cgg cac aac ctg   777
Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg His Asn Leu
        230                 235                 240

ggc ctg cag ctc tcg gtg gag acg ctg gat ggg cag agc atc aac ccc   825
Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser Ile Asn Pro
    245                 250                 255

aag ttg gcg ggc ctg att ggg cgg cac ggg ccc cag aac aag cag ccc   873
Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn Lys Gln Pro
260                 265                 270                 275

ttc atg gtg gct ttc ttc aag gcc acg gag gtc cac ttc cgc agc atc   921
Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Phe Arg Ser Ile
                280                 285                 290
```

```
cgg tcc acg ggg agc aaa cag cgc agc cag aac cgc tcc aag acg ccc   969
Arg Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr Pro
            295                 300                 305

aag aac cag gaa gcc ctg cgg atg gcc aac gtg gca gag aac agc agc   1017
Lys Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu Asn Ser Ser
        310                 315                 320

agc gac cag agg cag gcc tgt aag aag cac gag ctg tat gtc agc ttc   1065
Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe
    325                 330                 335

cga gac ctg ggc tgg cag gac tgg atc atc gcg cct gaa ggc tac gcc   1113
Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala
340                 345                 350                 355

gcc tac tac tgt gag ggg gag tgt gcc ttc cct ctg aac tcc tac atg   1161
Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met
                360                 365                 370

aac gcc acc aac cac gcc atc gtg cag acg ctg gtc cac ttc atc aac   1209
Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His Phe Ile Asn
            375                 380                 385

ccg gaa acg gtg ccc aag ccc tgc tgt gcg ccc acg cag ctc aat gcc   1257
Pro Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala
        390                 395                 400

atc tcc gtc ctc tac ttc gat gac agc tcc aac gtc atc ctg aag aaa   1305
Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile Leu Lys Lys
    405                 410                 415

tac aga aac atg gtg gtc cgg gcc tgt ggc tgc cac tagctcctcc        1351
Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
420                 425                 430

gagaattcag acccttttggg gccaagtttt tctggatcct ccattgctcg ccttggccag 1411

gaaccagcag accaactgcc ttttgtgaga ccttcccctc cctatcccca actttaaagg 1471

tgtgagagta ttaggaaaca tgagcagcat atggcttttg atcagttttt cagtggcagc 1531

atccaatgaa caagatccta caagctgtgc aggcaaaacc tagcaggaaa aaaaaacaac 1591

gcataaagaa aaatggccgg gccaggtcat tggctgggaa gtctcagcca tgcacggact 1651

cgtttccaga ggtaattatg agcgcctacc agccaggcca cccagccgtg ggaggaaggg 1711

ggcgtggcaa ggggtgggca cattggtgtc tgtgcgaaag gaaaattgac ccggaagttc 1771

ctgtaataaa tgtcacaata aaacgaatga atgaaaaaaa aaaaaaaaaa a          1822
```

<210> 39
<211> 431
<212> PRT
<213> Homo sapiens

<400> 39

```
Met His Val Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala
  1               5                  10                  15

Leu Trp Ala Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser
             20                  25                  30

Leu Asp Asn Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser
         35                  40                  45

Gln Glu Arg Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu
     50                  55                  60

Pro His Arg Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro
 65                  70                  75                  80

Met Phe Met Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Gly Gly
                 85                  90                  95

Gly Pro Gly Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser
            100                 105                 110

Thr Gln Gly Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr
        115                 120                 125

Asp Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys
    130                 135                 140

Glu Phe Phe His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu
145                 150                 155                 160

Ser Lys Ile Pro Glu Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile
                165                 170                 175

Tyr Lys Asp Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Arg Ile
            180                 185                 190

Ser Val Tyr Gln Val Leu Gln Glu His Leu Gly Arg Glu Ser Asp Leu
        195                 200                 205

Phe Leu Leu Asp Ser Arg Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu
    210                 215                 220

Val Phe Asp Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg
225                 230                 235                 240

His Asn Leu Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser
                245                 250                 255

Ile Asn Pro Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn
            260                 265                 270

Lys Gln Pro Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Phe
        275                 280                 285

Arg Ser Ile Arg Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser
```

```
    290                 295                 300

Lys Thr Pro Lys Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu
305                 310                 315                 320

Asn Ser Ser Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr
                325                 330                 335

Val Ser Phe Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu
            340                 345                 350

Gly Tyr Ala Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn
        355                 360                 365

Ser Tyr Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His
    370                 375                 380

Phe Ile Asn Pro Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln
385                 390                 395                 400

Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile
                405                 410                 415

Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
            420                 425                 430


<210> 40
<211> 98
<212> PRT
<213> Homo sapiens

<220>
<223> TGF-Beta1

<400> 40
Cys Cys Val Arg Gln Leu Tyr Ile Asp Phe Arg Lys Asp Leu Gly Trp
  1               5                  10                  15

Lys Trp Ile His Glu Pro Lys Gly Tyr His Ala Asn Phe Cys Leu Gly
             20                  25                  30

Pro Cys Pro Tyr Ile Trp Ser Leu Asp Thr Gln Tyr Ser Lys Val Leu
         35                  40                  45

Ala Leu Tyr Asn Gln His Asn Pro Gly Ala Ser Ala Ala Pro Cys Cys
     50                  55                  60

Val Pro Gln Ala Leu Glu Pro Leu Pro Ile Val Tyr Tyr Val Gly Arg
 65                  70                  75                  80

Lys Pro Lys Val Glu Gln Leu Ser Asn Met Ile Val Arg Ser Cys Lys
                 85                  90                  95

Cys Ser
```

<210> 41
<211> 98
<212> PRT
<213> Homo sapiens

<220>
<223> TGF-Beta2

<400> 41
Cys Cys Leu Arg Pro Leu Tyr Ile Asp Phe Lys Arg Asp Leu Gly Trp
1 5 10 15

Lys Trp Ile His Glu Pro Lys Gly Tyr Asn Ala Asn Phe Cys Ala Gly
20 25 30

Ala Cys Pro Tyr Leu Trp Ser Ser Asp Thr Gln His Ser Arg Val Leu
35 40 45

Ser Leu Tyr Asn Thr Ile Asn Pro Glu Ala Ser Ala Ser Pro Cys Cys
50 55 60

Val Ser Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Ile Gly Lys
65 70 75 80

Thr Pro Lys Ile Glu Gln Leu Ser Asn Met Ile Val Lys Ser Cys Lys
85 90 95

Cys Ser

<210> 42
<211> 98
<212> PRT
<213> Homo sapiens

<220>
<223> TGF-Beta3

<400> 42
Cys Cys Val Arg Pro Leu Tyr Ile Asp Phe Arg Gln Asp Leu Gly Trp
1 5 10 15

Lys Trp Val His Glu Pro Lys Gly Tyr Tyr Ala Asn Phe Cys Ser Gly
20 25 30

Pro Cys Pro Tyr Leu Arg Ser Ala Asp Thr Thr His Ser Thr Val Leu
35 40 45

Gly Leu Tyr Asn Thr Leu Asn Pro Glu Ala Ser Ala Ser Pro Cys Cys
50 55 60

Val Pro Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Val Gly Arg
65 70 75 80

Thr Pro Lys Val Glu Gln Leu Ser Asn Met Val Val Lys Ser Cys Lys

85 90 95

Cys Ser

<210> 43
<211> 98
<212> PRT
<213> Gallus gallus

<220>
<223> TGF-Beta4

<400> 43
Cys Cys Val Arg Pro Leu Tyr Ile Asp Phe Arg Lys Asp Leu Gln Trp
1 5 10 15

Lys Trp Ile His Glu Pro Lys Gly Tyr Met Ala Asn Phe Cys Met Gly
20 25 30

Pro Cys Pro Tyr Ile Trp Ser Ala Asp Thr Gln Tyr Thr Lys Val Leu
35 40 45

Ala Leu Tyr Asn Gln His Asn Pro Gly Ala Ser Ala Ala Pro Cys Cys
50 55 60

Val Pro Gln Thr Leu Asp Pro Leu Pro Ile Ile Tyr Tyr Val Gly Arg
65 70 75 80

Asn Val Arg Val Glu Gln Leu Ser Asn Met Val Val Arg Ala Cys Lys
85 90 95

Cys Ser

<210> 44
<211> 98
<212> PRT
<213> Xenopus laevis

<220>
<223> TGF-Beta5

<400> 44
Cys Cys Val Lys Pro Leu Tyr Ile Asn Phe Arg Lys Asp Leu Gly Trp
1 5 10 15

Lys Trp Ile His Glu Pro Lys Gly Tyr Glu Ala Asn Tyr Cys Leu Gly
20 25 30

Asn Cys Pro Tyr Ile Trp Ser Met Asp Thr Gln Tyr Ser Lys Val Leu
35 40 45

Ser Leu Tyr Asn Gln Asn Asn Pro Gly Ala Ser Ile Ser Pro Cys Cys
50 55 60

```
Val Pro Asp Val Leu Glu Pro Leu Pro Ile Ile Tyr Tyr Val Gly Arg
 65                  70                  75                  80

Thr Ala Lys Val Glu Gln Leu Ser Asn Met Val Val Arg Ser Cys Asn
                 85                  90                  95

Cys Ser
```

<210> 45
<211> 102
<212> PRT
<213> Drosophila melanogaster

<220>
<223> DPP

<400> 45

```
Cys Arg Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asp
  1               5                  10                  15

Asp Trp Ile Val Ala Pro Leu Gly Tyr Asp Ala Tyr Tyr Cys His Gly
             20                  25                  30

Lys Cys Pro Phe Pro Leu Ala Asp His Phe Asn Ser Thr Asn His Ala
         35                  40                  45

Val Val Gln Thr Leu Val Asn Asn Met Asn Pro Gly Lys Val Pro Lys
     50                  55                  60

Ala Cys Cys Val Pro Thr Gln Leu Asp Ser Val Ala Met Leu Tyr Leu
 65                  70                  75                  80

Asn Asp Gln Ser Thr Val Val Leu Lys Asn Tyr Gln Glu Met Thr Val
                 85                  90                  95

Val Gly Cys Gly Cys Arg
             100
```

<210> 46
<211> 102
<212> PRT
<213> Xenopus laevis

<220>
<223> VG1

<400> 46

```
Cys Lys Lys Arg His Leu Tyr Val Glu Phe Lys Asp Val Gly Trp Gln
  1               5                  10                  15

Asn Trp Val Ile Ala Pro Gln Gly Tyr Met Ala Asn Tyr Cys Tyr Gly
             20                  25                  30
```

```
Glu Cys Pro Tyr Pro Leu Thr Glu Ile Leu Asn Gly Ser Asn His Ala
         35                  40                  45

Ile Leu Gln Thr Leu Val His Ser Ile Glu Pro Glu Asp Ile Pro Leu
     50                  55                  60

Pro Cys Cys Val Pro Thr Lys Met Ser Pro Ile Ser Met Leu Phe Tyr
 65                  70                  75                  80

Asp Asn Asn Asp Asn Val Val Leu Arg His Tyr Glu Asn Met Ala Val
                 85                  90                  95

Asp Glu Cys Gly Cys Arg
            100


<210> 47
<211> 102
<212> PRT
<213> Mus musculus

<220>
<223> VGR1

<400> 47
Cys Lys Lys His Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Gln
  1               5                  10                  15

Asp Trp Ile Ile Ala Pro Lys Gly Tyr Ala Ala Asn Tyr Cys Asp Gly
             20                  25                  30

Glu Cys Ser Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His Ala
         35                  40                  45

Ile Val Gln Thr Leu Val His Leu Met Asn Pro Glu Tyr Val Pro Lys
     50                  55                  60

Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe
 65                  70                  75                  80

Asp Asp Asn Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val
                 85                  90                  95

Arg Ala Cys Gly Cys His
            100


<210> 48
<211> 118
<212> PRT
<213> Drosophila melanogaster

<220>
<223> 60A

<400> 48
Cys Gln Met Gln Thr Leu Tyr Ile Asp Phe Lys Asp Leu Gly Trp His
```

```
  1               5                  10                  15

Asp Trp Ile Ile Ala Pro Glu Gly Tyr Gly Ala Phe Tyr Cys Ser Gly
             20                  25                  30

Glu Cys Asn Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His Ala
         35                  40                  45

Ile Val Gln Thr Leu Val His Leu Leu Glu Pro Lys Lys Val Pro Lys
     50                  55                  60

Pro Cys Cys Ala Pro Thr Arg Leu Gly Ala Leu Pro Val Leu Tyr His
 65                  70                  75                  80

Pro Cys Cys Ala Pro Thr Arg Leu Gly Ala Leu Pro Val Leu Tyr His
                 85                  90                  95

Leu Asn Asp Glu Asn Val Asn Leu Lys Lys Tyr Arg Asn Met Ile Val
            100                 105                 110

Lys Ser Cys Gly Cys His
        115
```

<210> 49
<211> 101
<212> PRT
<213> Homo sapiens

<220>
<223> BMP-2A

<400> 49

```
Cys Lys Arg His Pro Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn
  1               5                  10                  15

Asp Trp Ile Val Ala Pro Pro Gly Tyr His Ala Phe Tyr Cys His Gly
             20                  25                  30

Glu Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr Asn His Ala
         35                  40                  45

Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Lys Ile Pro Lys Ala
     50                  55                  60

Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp
 65                  70                  75                  80

Glu Asn Glu Lys Val Val Leu Lys Asn Tyr Gln Asp Met Val Val Glu
                 85                  90                  95

Gly Cys Gly Cys Arg
            100
```

<210> 50
<211> 103

<212> PRT
<213> Homo sapiens

<220>
<223> BMP3

<400> 50
Cys Ala Arg Arg Tyr Leu Lys Val Asp Phe Ala Asp Ile Gly Trp Ser
1 5 10 15

Glu Trp Ile Ile Ser Pro Lys Ser Phe Asp Ala Tyr Tyr Cys Ser Gly
20 25 30

Ala Cys Gln Phe Pro Met Pro Lys Ser Leu Lys Pro Ser Asn His Ala
35 40 45

Thr Ile Gln Ser Ile Val Arg Ala Val Gly Val Val Pro Gly Ile Pro
50 55 60

Glu Pro Cys Cys Val Pro Glu Lys Met Ser Ser Leu Ser Ile Leu Phe
65 70 75 80

Phe Asp Glu Asn Lys Asn Val Val Leu Lys Val Tyr Pro Asn Met Thr
85 90 95

Val Glu Ser Cys Ala Cys Arg
100

<210> 51
<211> 101
<212> PRT
<213> Homo sapiens

<220>
<223> BMP-4

<400> 51
Cys Arg Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn
1 5 10 15

Asp Trp Ile Val Ala Pro Pro Gly Tyr Gln Ala Phe Tyr Cys His Gly
20 25 30

Asp Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr Asn His Ala
35 40 45

Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Ser Ile Pro Lys Ala
50 55 60

Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp
65 70 75 80

Glu Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu Met Val Val Glu
85 90 95

Gly Cys Gly Cys Arg

100

<210> 52
<211> 102
<212> PRT
<213> Homo sapiens

<220>
<223> BMP-5

<400> 52
Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp Gln
1 5 10 15

Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Phe Tyr Cys Asp Gly
20 25 30

Glu Cys Ser Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His Ala
35 40 45

Ile Val Gln Thr Leu Val His Leu Met Phe Pro Asp His Val Pro Lys
50 55 60

Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe
65 70 75 80

Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val
85 90 95

Arg Ser Cys Gly Cys His
100

<210> 53
<211> 102
<212> PRT
<213> Homo sapiens

<220>
<223> BMP-6

<400> 53
Cys Arg Lys His Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Gln
1 5 10 15

Asp Trp Ile Ile Ala Pro Lys Gly Tyr Ala Ala Asn Tyr Cys Asp Gly
20 25 30

Glu Cys Ser Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His Ala
35 40 45

Ile Val Gln Thr Leu Val His Leu Met Asn Pro Glu Tyr Val Pro Lys
50 55 60

Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe
65 70 75 80

```
Asp Asp Asn Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val
                 85                  90                  95

Arg Ala Cys Gly Cys His
            100


<210> 54
<211> 103
<212> PRT
<213> Gallus gallus

<220>
<223> DORSALIN

<400> 54
Cys Arg Arg Thr Ser Leu His Val Asn Phe Lys Glu Ile Gly Trp Asp
  1               5                  10                  15

Ser Trp Ile Ile Ala Pro Lys Asp Tyr Glu Ala Phe Glu Cys Lys Gly
             20                  25                  30

Gly Cys Phe Phe Pro Leu Thr Asp Asn Val Thr Pro Thr Lys His Ala
         35                  40                  45

Ile Val Gln Thr Leu Val His Leu Gln Asn Pro Lys Lys Ala Ser Lys
     50                  55                  60

Ala Cys Cys Val Pro Thr Lys Leu Asp Ala Ile Ser Ile Leu Tyr Lys
 65                  70                  75                  80

Asp Asp Ala Gly Val Pro Thr Leu Ile Tyr Asn Tyr Glu Gly Met Lys
                 85                  90                  95

Val Ala Glu Cys Gly Cys Arg
            100


<210> 55
<211> 102
<212> PRT
<213> Homo sapiens

<220>
<223> OP-1

<400> 55
Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp Gln
  1               5                  10                  15

Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Tyr Tyr Cys Glu Gly
             20                  25                  30

Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met Asn Ala Thr Asn His Ala
         35                  40                  45
```

```
Ile Val Gln Thr Leu Val His Phe Ile Asn Pro Glu Thr Val Pro Lys
     50                  55                  60

Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala Ile Ser Val Leu Tyr Phe
 65                  70                  75                  80

Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val
                 85                  90                  95

Arg Ala Cys Gly Cys His
            100


<210> 56
<211> 102
<212> PRT
<213> Homo sapiens

<220>
<223> OP-2

<400> 56
Cys Arg Arg His Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Leu
  1               5                  10                  15

Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys Glu Gly
             20                  25                  30

Glu Cys Ser Phe Pro Leu Asp Ser Cys Met Asn Ala Thr Asn His Ala
         35                  40                  45

Ile Leu Gln Ser Leu Val His Leu Met Lys Pro Asn Ala Val Pro Lys
     50                  55                  60

Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr Ser Val Leu Tyr Tyr
 65                  70                  75                  80

Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His Arg Asn Met Val Val
                 85                  90                  95

Lys Ala Cys Gly Cys His
            100


<210> 57
<211> 102
<212> PRT
<213> Mus musculus

<220>
<223> OP-3

<400> 57
Cys Arg Arg His Glu Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp Leu
  1               5                  10                  15

Asp Ser Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr Cys Ala Gly
```

```
            20                  25                  30

Glu Cys Ile Tyr Pro Leu Asn Ser Cys Met Asn Ser Thr Asn His Ala
         35                  40                  45

Thr Met Gln Ala Leu Val His Leu Met Lys Pro Asp Ile Ile Pro Lys
     50                  55                  60

Val Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser Leu Leu Tyr Tyr
 65                  70                  75                  80

Asp Arg Asn Asn Asn Val Ile Leu Arg Arg Glu Arg Asn Met Val Val
                 85                  90                  95

Gln Ala Cys Gly Cys His
            100


<210> 58
<211> 107
<212> PRT
<213> Mus musculus

<220>
<223> GDF-1

<400> 58
Cys Arg Thr Arg Arg Leu His Val Ser Phe Arg Glu Val Gly Trp His
  1               5                  10                  15

Arg Trp Val Ile Ala Pro Arg Gly Phe Leu Ala Asn Phe Cys Gln Gly
             20                  25                  30

Thr Cys Ala Leu Pro Glu Thr Leu Arg Gly Pro Gly Gly Pro Pro Ala
         35                  40                  45

Leu Asn His Ala Val Leu Arg Ala Leu Met His Ala Ala Ala Pro Thr
     50                  55                  60

Pro Gly Ala Gly Ser Pro Cys Cys Val Pro Glu Arg Leu Ser Pro Ile
 65                  70                  75                  80

Ser Val Leu Phe Phe Asp Asn Ser Asp Asn Val Val Leu Arg His Tyr
                 85                  90                  95

Glu Asp Met Val Val Asp Glu Cys Gly Cys Arg
            100                 105


<210> 59
<211> 101
<212> PRT
<213> Mus musculus

<220>
<223> GDF-3
```

<400> 59
Cys His Arg His Gln Leu Phe Ile Asn Phe Gln Asp Leu Gly Trp His
1 5 10 15

Lys Trp Val Ile Ala Pro Lys Gly Phe Met Ala Asn Tyr Cys His Gly
20 25 30

Glu Cys Pro Phe Ser Met Thr Thr Tyr Leu Asn Ser Ser Asn Tyr Ala
35 40 45

Phe Met Gln Ala Leu Met His Met Ala Asp Pro Lys Val Pro Lys Ala
50 55 60

Val Cys Val Pro Thr Lys Leu Ser Pro Ile Ser Met Leu Tyr Gln Asp
65 70 75 80

Ser Asp Lys Asn Val Ile Leu Arg His Tyr Glu Asp Met Val Val Asp
85 90 95

Glu Cys Gly Cys Gly
100

<210> 60
<211> 102
<212> PRT
<213> Mus musculus

<220>
<223> GDF-9

<400> 60
Cys Glu Leu His Asp Phe Arg Leu Ser Phe Ser Gln Leu Lys Trp Asp
1 5 10 15

Asn Trp Ile Val Ala Pro His Arg Tyr Asn Pro Arg Tyr Cys Lys Gly
20 25 30

Asp Cys Pro Arg Ala Val Arg His Arg Tyr Gly Ser Pro Val His Thr
35 40 45

Met Val Gln Asn Ile Ile Tyr Glu Lys Leu Asp Pro Ser Val Pro Arg
50 55 60

Pro Ser Cys Val Pro Gly Lys Tyr Ser Pro Leu Ser Val Leu Thr Ile
65 70 75 80

Glu Pro Asp Gly Ser Ile Ala Tyr Lys Glu Tyr Glu Asp Met Ile Ala
85 90 95

Thr Arg Cys Thr Cys Arg
100

<210> 61
<211> 105
<212> PRT

<213> Homo sapiens

<220>
<223> INHIBIN-Alpha

<400> 61
Cys His Arg Val Ala Leu Asn Ile Ser Phe Gln Glu Leu Gly Trp Glu
1 5 10 15

Arg Trp Ile Val Tyr Pro Pro Ser Phe Ile Phe His Tyr Cys His Gly
20 25 30

Gly Cys Gly Leu His Ile Pro Pro Asn Leu Ser Leu Pro Val Pro Gly
35 40 45

Ala Pro Pro Thr Pro Ala Gln Pro Tyr Ser Leu Leu Pro Gly Ala Gln
50 55 60

Pro Cys Cys Ala Ala Leu Pro Gly Thr Met Arg Pro Leu His Val Arg
65 70 75 80

Thr Thr Ser Asp Gly Gly Tyr Ser Phe Lys Tyr Glu Thr Val Pro Asn
85 90 95

Leu Leu Thr Gln His Cys Ala Cys Ile
100 105

<210> 62
<211> 106
<212> PRT
<213> Bos taurus

<220>
<223> INHIBIN-BetaA

<400> 62
Cys Cys Lys Lys Gln Phe Phe Val Ser Phe Lys Asp Ile Gly Trp Asn
1 5 10 15

Asp Trp Ile Ile Ala Pro Ser Gly Tyr His Ala Asn Tyr Cys Glu Gly
20 25 30

Glu Cys Pro Ser His Ile Ala Gly Thr Ser Gly Ser Ser Leu Ser Phe
35 40 45

His Ser Thr Val Ile Asn His Tyr Arg Met Arg Gly His Ser Pro Phe
50 55 60

Ala Asn Leu Lys Ser Cys Cys Val Pro Thr Lys Leu Arg Pro Met Ser
65 70 75 80

Met Leu Tyr Tyr Asp Asp Gly Gln Asn Ile Ile Lys Lys Asp Ile Gln
85 90 95

Asn Met Ile Val Glu Glu Cys Gly Cys Ser
100 105

<210> 63
<211> 106
<212> PRT
<213> Homo sapiens

<220>
<223> INHIBIN-BetaB

<400> 63
Cys Cys Lys Lys Gln Phe Phe Val Ser Phe Lys Asp Ile Gly Trp Asn
1 5 10 15

Asp Trp Ile Ile Ala Pro Ser Gly Tyr His Ala Asn Tyr Cys Glu Gly
20 25 30

Glu Cys Pro Ser His Ile Ala Gly Thr Ser Gly Ser Ser Leu Ser Phe
35 40 45

His Ser Thr Val Ile Asn His Tyr Arg Met Arg Gly His Ser Pro Phe
50 55 60

Ala Asn Leu Lys Ser Cys Cys Val Pro Thr Lys Leu Arg Pro Met Ser
65 70 75 80

Met Leu Tyr Tyr Asp Asp Gly Gln Asn Ile Ile Lys Lys Asp Ile Gln
85 90 95

Asn Met Ile Val Glu Glu Cys Gly Cys Ser
100 105

<210> 64
<211> 98
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: TGF-B
SUBGROUP SEQUENCE PATTERN

<220>
<223> Each Xaa is independently selected from a group of
one or more specified amino acids as defined in
the specification

<400> 64
Cys Cys Val Arg Pro Leu Tyr Ile Asp Phe Arg Xaa Asp Leu Gly Trp
1 5 10 15

Lys Trp Ile His Glu Pro Lys Gly Tyr Xaa Ala Asn Phe Cys Xaa Gly
20 25 30

Xaa Cys Pro Tyr Xaa Trp Ser Xaa Asp Thr Gln Xaa Ser Xaa Val Leu
35 40 45

Xaa Leu Tyr Asn Xaa Xaa Asn Pro Xaa Ala Ser Ala Xaa Pro Cys Cys
50 55 60

Val Pro Gln Xaa Leu Glu Pro Leu Xaa Ile Xaa Tyr Tyr Val Gly Arg
65 70 75 80

Xaa Xaa Lys Val Glu Gln Leu Ser Asn Met Xaa Val Xaa Ser Cys Lys
85 90 95

Cys Ser

<210> 65
<211> 104
<212> PRT
<213> Artificial Sequence

<220>
<223> Each Xaa is independently selected from a group of one or more specified amino acids as defined in the specification

<220>
<223> Description of Artificial Sequence: VG/DPP SUBGROUP SEQUENCE PATTERN

<400> 65
Cys Xaa Xaa Xaa Xaa Leu Tyr Val Xaa Phe Xaa Asp Xaa Gly Trp Xaa
1 5 10 15

Asp Trp Ile Ile Ala Pro Xaa Gly Tyr Xaa Ala Xaa Tyr Cys Xaa Gly
20 25 30

Xaa Cys Xaa Phe Pro Leu Xaa Xaa Xaa Xaa Asn Xaa Thr Asn His Ala
35 40 45

Ile Xaa Gln Thr Leu Val Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro
50 55 60

Lys Xaa Cys Cys Xaa Pro Thr Xaa Leu Xaa Ala Xaa Ser Xaa Leu Tyr
65 70 75 80

Xaa Asp Xaa Xaa Xaa Xaa Xaa Val Xaa Leu Xaa Xaa Tyr Xaa Xaa Met
85 90 95

Xaa Val Xaa Xaa Cys Gly Cys Xaa
100

<210> 66
<211> 107
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: GDF SUBGROUP

PATTERN

<220>
<223> Each Xaa is independently selected from a group of
one or more specified amino acids as defined in
the specification

<400> 66
Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Trp Xaa
1 5 10 15

Xaa Trp Xaa Xaa Ala Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Xaa Gly
20 25 30

Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
50 55 60

Pro Xaa Xaa Xaa Xaa Xaa Xaa Cys Val Pro Xaa Xaa Xaa Ser Pro Xaa
65 70 75 80

Ser Xaa Leu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Tyr
85 90 95

Glu Asp Met Xaa Xaa Xaa Xaa Cys Xaa Cys Xaa
100 105

<210> 67
<211> 109
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: INHIBIN
SUBGROUP PATTERN

<220>
<223> Each Xaa is independently selected from a group of
one or more specified amino acids as defined in
the specification

<400> 67
Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Gly Trp Xaa
1 5 10 15

Xaa Trp Ile Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Tyr Cys Xaa Gly
20 25 30

Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
35 40 45

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
50 55 60

```
Xaa Xaa Xaa Xaa Xaa Cys Cys Xaa Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa
 65                  70                  75                  80

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asp Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
                 85                  90                  95

Xaa Xaa Xaa Asn Xaa Xaa Xaa Xaa Xaa Cys Xaa Cys Xaa
            100                 105
```

<210> 68
<211> 139
<212> PRT
<213> Homo sapiens

<220>
<223> Mature H2223 mutant

<400> 68

```
Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr Pro Lys
  1               5                  10                  15

Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu Asn Ser Ser Ser
             20                  25                  30

Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg
         35                  40                  45

Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala
     50                  55                  60

Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met Asn
 65                  70                  75                  80

Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His Phe Ile Asn Pro
                 85                  90                  95

Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala Ile
            100                 105                 110

Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr
        115                 120                 125

Glu Asp Met Val Val Glu Ala Cys Gly Cys Arg
    130                 135
```

<210> 69
<211> 117
<212> PRT
<213> Homo sapiens

<220>
<223> Trypsin truncated H2223 mutant

<400> 69

```
Met Ala Asn Val Ala Glu Asn Ser Ser Ser Asp Gln Arg Gln Ala Cys
```

```
  1               5                  10                  15

Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp Gln Asp
            20                  25                  30

Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Tyr Tyr Cys Glu Gly Glu
        35                  40                  45

Cys Ala Phe Pro Leu Asn Ser Tyr Met Asn Ala Thr Asn His Ala Ile
    50                  55                  60

Val Gln Thr Leu Val His Phe Ile Asn Pro Glu Thr Val Pro Lys Pro
65                  70                  75                  80

Cys Cys Ala Pro Thr Gln Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp
                85                  90                  95

Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Glu Asp Met Val Val Glu
            100                 105                 110

Ala Cys Gly Cys Arg
        115
```

<210> 70
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer #1

<220>
<221> CDS
<222> (1)..(33)

<400> 70
gcg ccc acg cag ctc agc gct atc tcc gtc ctc 33
Ala Pro Thr Gln Leu Ser Ala Ile Ser Val Leu
  1   5   10

<210> 71
<211> 11
<212> PRT
<213> Artificial Sequence

<400> 71
Ala Pro Thr Gln Leu Ser Ala Ile Ser Val Leu
  1   5   10

<210> 72
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer #2

<400> 72
ctatctgcag ccacaagctt cgaccaccat gtcttcgtat ttc 43

<210> 73
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:complement of
Primer #2

<220>
<221> CDS
<222> (2)..(43)

<400> 73
g aaa tac gaa gac atg gtg gtc gaa gct tgt ggc tgc aga tag 43
Lys Tyr Glu Asp Met Val Val Glu Ala Cys Gly Cys Arg
1 5 10

<210> 74
<211> 13
<212> PRT
<213> Artificial Sequence

<400> 74
Lys Tyr Glu Asp Met Val Val Glu Ala Cys Gly Cys Arg
1 5 10

<210> 75
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:the sequence
between the T7 promoter, at the XbaI site, and the
ATG codon

<400> 75
tctagaataa ttttgtttaa cctttaagaa ggagatatac gatg 44

<210> 76
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer #3

<400> 76
taatacgact cactatagg 19

<210> 77
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer #4

<400> 77
gctgagctgc gtgggcgc 18

<210> 78
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: complement of
Primer #4

<220>
<221> CDS
<222> (1)..(18)

<400> 78
gcg ccc acg cag ctc agc 18
Ala Pro Thr Gln Leu Ser
1 5

<210> 79
<211> 6
<212> PRT
<213> Artificial Sequence

<400> 79
Ala Pro Thr Gln Leu Ser
1 5

<210> 80
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer #5

<400> 80
ggatcctatc tgcagccaca agc 23

<210> 81
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: complement of
Primer #5

<220>
<221> CDS
<222> (1)..(18)

<400> 81
gct tgt ggc tgc aga tag gatcc 23
Ala Cys Gly Cys Arg
1 5

<210> 82
<211> 5
<212> PRT
<213> Artificial Sequence

<400> 82
Ala Cys Gly Cys Arg
1 5

<210> 83
<211> 102
<212> PRT
<213> Homo sapiens

<220>
<223> CDMP-1/GDF-5

<400> 83
Cys Ser Arg Lys Ala Leu His Val Asn Phe Lys Asp Met Gly Trp Asp
1 5 10 15

Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His Cys Glu Gly
20 25 30

Leu Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala
35 40 45

Val Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser Thr Pro Pro
50 55 60

Thr Cys Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile Leu Phe Ile
65 70 75 80

Asp Ser Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val
85 90 95

Glu Ser Cys Gly Cys Arg
100

<210> 84
<211> 102
<212> PRT
<213> Homo sapiens

<220>
<223> CDMP-2/GDF-6

<400> 84
Cys Ser Lys Lys Pro Leu His Val Asn Phe Lys Glu Leu Gly Trp Asp
1 5 10 15

Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu Gly
20 25 30

Val Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala
35 40 45

Ile Ile Gln Thr Leu Met Asn Ser Met Asp Pro Gly Ser Thr Pro Pro
50 55 60

Ser Cys Cys Val Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile
65 70 75 80

Asp Ala Gly Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val
85 90 95

Glu Ser Cys Gly Cys Arg
100

<210> 85
<211> 102
<212> PRT
<213> Mus musculus

<220>
<223> GDF-6

<400> 85
Cys Ser Arg Lys Pro Leu His Val Asn Phe Lys Glu Leu Gly Trp Asp
1 5 10 15

Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu Gly
20 25 30

Val Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala
35 40 45

Ile Ile Gln Thr Leu Met Asn Ser Met Asp Pro Gly Ser Thr Pro Pro
50 55 60

Ser Cys Cys Val Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile

65 70 75 80

Asp Ala Gly Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val
85 90 95

Glu Ser Cys Gly Cys Arg
100

<210> 86
<211> 102
<212> PRT
<213> Bos taurus

<220>
<223> CDMP-2

<400> 86
Cys Ser Lys Lys Pro Leu His Val Asn Phe Lys Glu Leu Gly Trp Asp
1 5 10 15

Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu Gly
20 25 30

Val Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala
35 40 45

Ile Ile Gln Thr Leu Met Asn Ser Met Asp Pro Gly Ser Thr Pro Pro
50 55 60

Ser Cys Cys Val Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile
65 70 75 80

Asp Ala Gly Asn Asn Val Val Tyr Asn Glu Tyr Glu Glu Met Val Val
85 90 95

Glu Ser Cys Gly Cys Arg
100

<210> 87
<211> 102
<212> PRT
<213> Mus musculus

<220>
<223> GDF-7

<400> 87
Cys Ser Arg Lys Ser Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp
1 5 10 15

Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly
20 25 30

Val Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala
35 40 45

```
Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala
     50                  55                  60

Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile
 65                  70                  75                  80

Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val
                 85                  90                  95

Glu Ala Cys Gly Cys Arg
            100
```

<210> 88
<211> 102
<212> PRT
<213> Homo sapiens

<220>
<223> CDMP-3 construct

<400> 88

```
Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp
  1               5                  10                  15

Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly
             20                  25                  30

Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala
         35                  40                  45

Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala
     50                  55                  60

Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile
 65                  70                  75                  80

Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val
                 85                  90                  95

Glu Ala Cys Gly Cys Arg
            100
```